Yersinia enterocolitica specific infection by bacteriophages TG1 and φR1-RT is dependent on temperature regulated expression of the phage host receptor OmpF

Carlos G. Leon-Velarde,[a,f] Lotta Happonen,[i,k] Maria Pajunen,[g] Katarzyna Leskinen,[g] Andrew M. Kropinski,[b,c] Laura Mattinen,[g] Monika Rajtor,[g] Joanna Zur,[g] Darren Smith,[j] Shu Chen,[a] Ayesha Nawaz,[g] Roger P. Johnson,[d] Joseph A. Odumeru,[f] Mansel W. Griffiths,[e,f] Mikael Skurnik,[g,h]#

Laboratory Services Division, University of Guelph, Guelph, Ontario, Canada[a]; Department of Pathobiology, University of Guelph, Guelph, Ontario, Canada[b]; Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario, Canada[c]; National Microbiology Laboratory at Guelph, Public Health Agency of Canada, Guelph Ontario, Canada[d]; Canadian Research Institute for Food Safety, University of Guelph, Guelph, Ontario, Canada[e]; Department of Food Science, University of Guelph, Guelph, Ontario, Canada[f]; Department of Bacteriology and Immunology, Medicum, and Research Programs Unit, Immunobiology, University of Helsinki, Helsinki, Finland[g]; Division of Clinical Microbiology, Helsinki University Hospital, HUSLAB, Helsinki, Finland[h]; Department of Clinical Sciences Lund, Infection Medicine, Lund University, Lund, Sweden[i]; Applied Sciences, University of Northumbria, Newcastle upon Tyne, United Kingdom[j]; Institute of Biotechnology and Department of Biosciences, University of Helsinki, Helsinki, Finland[k]

Running Head: *Y. enterocolitica* bacteriophages TG1 and φR1-RT

Keywords: *Yersinia enterocolitica*, bacteriophage, receptor, OmpF, lipopolysaccharide, genome

#Address correspondence to Mikael Skurnik, mikael.skurnik@helsinki.fi.

ABSTRACT

Bacteriophages present huge potential both as a resource for developing novel tools for bacterial diagnostics and for use in phage therapy. This is also valid for bacteriophages specific for *Yersinia enterocolitica.* To increase our knowledge on *Y. enterocolitica* –specific phages we characterized two novel yersiniophages. The genomes of the bacteriophages vB_YenM_TG1 (TG1) and vB_YenM_ϕR1-RT (ϕR1-RT), isolated from pig manure in Canada and from sewage in Finland, consist of linear double-stranded DNA of 162,101 and 168,809 bp respectively. Their genomes encode 262 putative coding sequences and 4 tRNAs genes, and share 91% overall nucleotide identity. Based on phylogenetic analyses of their whole genome sequences and large terminase subunit protein sequences, a genus named *Tg1virus* within the family *Myoviridae* is proposed with TG1 and ϕR1-RT as member species. These bacteriophages exhibit a host range restricted to *Y. enterocolitica*, and display lytic activity against the epidemiologically significant serotypes O:3, O:5,27, and O:9 at and below 25°C. Adsorption analyses of LPS and OmpF mutants demonstrate that these phages use both the LPS inner core heptosyl residues and the outer membrane protein OmpF as phage receptors. Based on RNA-sequencing and quantitative proteomics we also demonstrate the temperature dependent infection is due to strong repression of OmpF at 37°C. In addition, ϕR1-RT was shown to be able to enter into a pseudolysogenic state. All together, this work provides further insight into phage-host cell interactions by highlighting the importance of understanding underlying factors which may affect the abundance of phage host receptors on the cell surface.

IMPORTANCE

Only a small number of bacteriophages infecting *Y. enterocolitica*, the predominant causative agent of yersiniosis, have been previously described. Here, two newly isolated *Y. enterocolitica* phages were studied in detail with the aim of elucidating the host cell receptors required for infection. Our research further expands the repertoire of phages available for consideration as potential antimicrobial agents or as diagnostic tools for this important bacterial pathogen.

INTRODUCTION

Yersinia enterocolitica, a facultative anaerobic, Gram-negative, non-sporulating, short bacillus isolated frequently from soil, water, animals, and foods, is an important zoonotic pathogen leading to human and animal enteric infection (1). The main animal reservoir for *Y. enterocolitica* is pigs, and pork derived products are thought to be the main source of human infections in addition to drinking of contaminated water and blood-transfusions (1, 2). Symptoms of yersiniosis may include diarrhea, terminal ileitis, mesenteric lymphadenitis, and septicemia (3). Among the species within the genus *Yersinia, Y. enterocolitica* is highly heterogeneous and is grouped into six phylogroups (4). The widely used bioserotype groups form the basis of the phylogroups such that phylogroup 1 contains the biotype 1A strains, phylogroup 2 the highly pathogenic biotype 1B strains, phylogroup 3 the bioserotype 4/O:3 strains, phylogroup 4 bioserotype 3/O:9 strains, phylogroup 5 bioserotype 2/O:5,27 strains and phylogroup 6 the serotype O:2,3 strains rarely isolated from hares (4–7). *Y. enterocolitica* is also represented by over 60 serotypes that are determined by the variability of O-antigens present in the outer cell membrane (8, 9). The predominant pathogenic strains associated with yersiniosis belong to bioserotypes 1B/O:8, 2/O:5,27, 2/O:9, 3/O:3, and 4/O:3, with the last being the most common in Europe, Japan, Canada, and the United States (1, 2). From 2010–2012, 98% of all reported yersiniosis infections worldwide were acquired in Europe, and most (97%) were caused by *Y. enterocolitica,* with the remainder caused by *Y. pseudotuberculosis* (10). In 2015, the most commonly reported *Y. enterocolitica* serotype in the European Union was O:3 (89%), followed by serotypes O:9 (7%), O:5,27 (2%) and O:8 (2%) (10).

Although several bacteriophages infecting *Y. enterocolitica* have been described, few have been studied in detail providing reliable information on morphology, host range, and or receptor specificity. To date, bacteriophages ϕYeO3-12(11–13) and vB_YenP_AP5 (14) with specificity for *Y. enterocolitica* O:3, phage PY54 exhibiting a host range restricted to *Y. enterocolitica* O:5 and O:5,27 (15), *Yersinia* phage ϕR1-37 with a broad host range within the species *Y. enterocolitica* (16, 17) and *Yersinia* phage PY-100 (18) exhibiting a broader host range restricted to the genus *Yersinia,* have been described. These bacteriophages use different parts of the *Y. enterocolitica* lipopolysaccharide (LPS) as receptors (19). Analysis of the host-range combined with genetic and structural data have shown that the receptor for ϕR1-37 is the *Y. enterocolitica* O:3 LPS outer core (OC) hexasaccharide (16). The host receptor for phages ϕYeO3-12 and vB_YenP_AP5 has been determined to be the LPS O-antigen of serotype O:3 consisting of the sugar 6-deoxy-L-altropyranose (12, 14, 20). Given the interest in bacteriophages because of their potential use as therapeutic, diagnostic, and bio-control agents, the aim of this study was to characterize two newly isolated bacteriophages that are active against several epidemiologically significant *Y. enterocolitica* serotypes. In this study, the genome characterization, morphology, host range, host cell receptor specificity, and taxonomic position of the myovirus phages vB_YenM_TG1 (hereafter TG1) and vB_YenM_ϕR1-RT (hereafter ϕR1-RT) are described.

MATERIALS AND METHODS

Bacterial strains, phage isolation, and growth conditions. Bacterial strains, bacteriophages and plasmids are listed in **Table 1**. Bacteriophage ϕR1-RT was isolated from the incoming sewage of the Turku (Finland) city sewage treatment

plant, as described for other viruses (19) whereas bacteriophage TG1 was isolated from pig manure collected from a rural farm in Ontario, Canada as described previously for the isolation of *Y. enterocolitica* phages for phagetyping (21). For DNA extraction and morphological studies, ϕR1-RT was propagated on *Y. enterocolitica* strain YeO3-R1 (22) and TG1 on *Y. enterocolitica* strain YeO3-c (23).

Electron Microscopy. The preparation of the phage particles for transmission electron microscopy (TEM) was done as described (17, 24). Details are presented in Supplementary Materials and Methods.

Host Range. The lytic activity of ϕR1-RT and TG1 was tested on 109 and 160 strains (**Table S1**), respectively, belonging to 13 *Yersinia* species, as determined by standard spot tests (24). Briefly, 10 µl from a phage suspension containing approximately 10^8 PFU were spotted in the middle of a lawn of bacteria incubating for 18-24 h. Each strain was tested three times at 25°C and at 37 °C. Bacterial strains were considered sensitive to the phage if the degree of lysis was observed as a complete clearing, clearing throughout but with a faint hazy background, substantial turbidity throughout the cleared zone, or a few individual plaques (24). Bacterial strains were considered resistant if there was no effect of the phage on bacterial growth.

Genome sequencing and assembly. Details of the determination of the genomic sequences of phages ϕR1-RT and TG1 as well as the draft genomes of *Y. enterocolitica* strains YeO3-ϕR1-RT-R2, –R7 and –R9 are presented in Supplementary Materials and Methods.

Bioinformatics. Detailed description of the bioinformatics tools used is given in Supplementary Materials and Methods.

Complementation of the *Y. enterocolitica* O:3 OmpF mutant. The full ORF of *ompF* gene plus the upstream promoter region of YeO3-c was cloned as a 2 kb PCR fragment that was amplified with Phusion DNA polymerase using primer pair OmpC-F2 and OmpC-R2 (**Table S2**) into plasmids pTM100 and pSW25T to obtain plasmids pTM100_OmpF and pSW25T_OmpF, respectively (**Table 1**). Briefly, the PCR fragments were digested with MfeI and ligated with EcoRI digested, SAP-treated pTM100 or pSW25T. The constructed plasmids were mobilized to the OmpF mutant strain YeO3-c-R1-Cat17 by diparental conjugation as described earlier (25).

Phage adsorption assay. To identify the phage cell host receptors, a variety of *Y. enterocolitica* O:3 mutants (**Table 1**) were utilized in phage adsorption experiments. Approximately $5x10^3$ PFU of phage ϕR1-RT or phage TG1 in 100 μl was mixed with a 400-μl sample of bacteria (A_{600} ~1.2). The suspension was incubated at RT for 5 min and centrifuged at 16,000 *g* for 3 min, and the phage titer remaining in the supernatant, *i.e.,* the residual PFU percentage, was determined. LB was used as a non-adsorbing control in each assay, and the phage titer in the control supernatant was set to 100%. Each assay was performed in duplicate and repeated at least three times.

Total RNA extraction and RNA sequencing. Detailed description of the methods is presented in Supplementary information Materials and Methods. The RNA sequence data has been deposited to Gene Expression Omnibus (accession number GSE66516).

Quantitative proteomics. Detailed description of the methods is presented in Supplementary information Materials and Methods.

Transduction assay. *Y. enterocolitica* O:3 strain YeO3-*hfq::*Km with *hfq* gene knocked-out with a kanamycin resistance cassette (**Table 1**) was used as a donor and transducing particles were produced by infecting this strain with phage ϕR1-RT using the soft agar overlay method. Following overnight incubation, phages were eluted from the soft agar using SM buffer. The transducing lysates were centrifuged and treated with chloroform to prevent contamination with the donor strain. The titer of the obtained transducing stock was 6.62 x 10^9 PFU/mL. The *Y. enterocolitica* strain 6471/76 was used as the recipient. For the transduction of the recipient strain, 10 one mL aliquots of log-phase bacterial cultures containing 10^9 CFU/mL cells were mixed with 100 µl of 10^{-2} diluted transducing phage stock resulting at MOI of 0.006. After 15 min the bacterial cells were centrifuged and washed with LB and centrifuging them down removed the unabsorbed phages. The final cell pellet was resuspended in 100 µl LB, and the cells were allowed to recover during 30 min incubation with vigorous shaking. Subsequently, the bacterial cultures were plated on urea agar plates (0.1% peptone, 0.1% glucose, 0.5% NaCl, 0.2% KH_2PO_4, 0.00012% phenol red, 2% urea, 1.5% agar) supplemented with kanamycin (200 µg/mL) and incubated for 48h. The kanamycin resistant and urease negative colonies were considered as transduced. The transducing stock was also plated to ensure no contamination with donor strain.

Growth curves. Overnight bacterial cultures were diluted 1:10 in fresh LB medium and 180 µl aliquots were distributed into honeycomb plate wells (Growth Curves Ab Ltd) where they were mixed with 20 µl aliquots of different ϕR1-RT phage stock

184 dilutions (10^0 – 10^{-4}). A negative control was obtained by mixing 20 µl of phage stock
185 with 180 µl of medium, whereas positive control consisted of 180 µl of bacterial
186 culture and 20 µl of medium. The growth experiments were carried out at 4°C, 10°C,
187 16°C, 22°C, and 37°C using the Bioscreen C incubator (Growth Curves Ab Ltd) with
188 continuous shaking. The OD_{600} of the cultures was measured at selected time
189 intervals. The averages were calculated from values obtained for the bacteria grown in
190 5 parallel wells.
191
192 **Phage resistant mutant isolation**. A culture of wild type *Y. enterocolitica* strain
193 6471/76 was used to flood LB agar plates (LA). After the excess fluid was removed
194 the plates were allowed to dry before two 100 µl aliquots of the ϕR1-RT stock were
195 pipetted on the lawn of cells. The plates were incubated at 22°C and inspected daily
196 for phage resistant colonies growing within the lysis zones. After three days several
197 colonies appeared and among them three confirmed phage resistant derivatives were
198 isolated. The strains were named YeO3-ϕR1-RT-R2, YeO3-ϕR1-RT-R7, and YeO3-
199 ϕR1-RT-R9.
200
201 **CatMu-library screening**. The *CatMu*-transposon insertion library in *Y.*
202 *enterocolitica* strain YeO3-R1 has been described elsewhere (26, 27). In the present
203 work, a library representing 16,000 independent insertion mutants was screened. The
204 library was grown in LA supplemented with 100 µg/ml chloramphenicol (LA-Clm)
205 until OD_{600} = ~0.5. Phage ϕR1-RT was added to 1 mL of the library culture at MOI
206 ~10, fresh LB added to 5 mL and the culture was incubated at 22°C for 2h during
207 which time all phage-sensitive bacteria were expected to be infected and lysed. The
208 surviving bacteria were pelleted by centrifugation, washed twice with 1 mL LB to

remove remaining phages and after resuspending into 100 µl of LB plated on four LA-Clm plates that were incubated at 22°C. The Clm^R colonies were re-streaked on LA-Clm plates for further study.

Arbitrary PCR. Detailed description of the method is presented in Supplementary information Materials and Methods.

Cloning, expression and purification of the phage long tail fiber host receptor binding protein. The phage TG1 distal long tail fiber (LTF) protein Gp37 was co-expressed with phage encoded chaperones Gp57A and Gp38 to synthesize the native form of the putative receptor binding protein (RBP) as describer previously for the LTF of phage T4 (28). The Gp37 encoding gene was first cloned into the multiple cloning site (MCS) 1 of pCDF Duet-1 (conferring streptomycin resistance), producing pCDF Duet-1 Gp37. Then, the Gp38 encoding gene was cloned into the MCS 2 of pCDF Duet-1 Gp37, yielding pCDF Duet-1 Gp37-Gp38. Plasmid pET21a(+) conferring ampicillin resistance was used to clone the chaperone Gp57A encoding gene yielding plasmid pET21a(+) Gp57A. The plasmid constructs carry under the control of promoter T7, high level inducible gene expression with a His_6 fusion tag at the N-terminus for purification by chelating affinity chromatography (**Fig. S1**). The genes encoding Gp38 and Gp57A however, were expressed without a purification tag. PCR, restriction analysis, and DNA sequencing were used to verify the structure of the plasmids. For expression, *E. coli* BL21 Star™ (DE3) PLysS cells (Invitrogen) were transformed with pCDF-Duet-1 Gp37 or pCDF-Duet-1 Gp37-Gp38 and the same plasmids were also co-transformed with pET21a(+) Gp57A. Plasmid bearing *E. coli* were grown aerobically at 37°C to an OD_{600} = ~0.6 with shaking at 200 rpm in

250 mL of 2xYT media (16 g/L tryptone, 10 g/L yeast extract, 5.0 g/L NaCl, 0.22 μm filter sterilized, pH 6.5-7.5) supplemented with 50 μg/mL of ampicillin and or 50 μg/mL streptomycin as required. Protein expression was induced by the addition of 1 mM isopropyl-d-1-thiogalactopyranoside (IPTG) (Sigma-Aldrich, USA) incubating for 24h at 30°C with shaking at 200 rpm. Cells were harvested by centrifugation at 10,000 *g* for 15 min at 4°C and the pellets were resuspended in 25 mL of buffer A (50 mM sodium phosphate, 300 mM NaCl, 10mM imidazole, pH 8.0) supplemented with a protease inhibitor cocktail (Roche). Cells were disrupted by 10 rounds of 15 s of sonication using a Virsonic Digital 475 ultrasonicator (VirTis, NY, USA) alternating with incubation on ice. Insoluble debris was removed by centrifugation at 18,000 *g* for 30 min at 4°C and the soluble fraction was filtered through a 0.22 μm pore size filter (EMD Millipore, USA). The protein was purified by immobilized metal ion affinity chromatography using a nickel-nitrilotriacetic acid (Ni-NTA) agarose column (Novex, Invitrogen) according to the manufacturer's protocol. Captured proteins were eluted from the column using buffer B (50 mM sodium phosphate, 300 mM NaCl, 500mM imidazole, pH 8.0) and concentrated using Amicon-Pro centrifuge filters (Millipore) with a 10,000 Da molecular mass exclusion limit incorporating three washes with 10 mM Tris–HCl of pH 8.5. Protein concentration was estimated by measuring sample absorbance at 280 and 260 nm using a Nanodrop 2000 UV-vis Spectrophotometer (Thermo Scientific, USA) and Qubit® Protein Assay Kit using a Qubit® 1.0 fluorometer (Life Technologies) as per the manufacturer instructions. Protein analysis was performed by Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (29) using Mini-Protean®TGX Stain-Free Precast Gels (Bio-Rad Laboratories, USA) and Coomassie blue staining. Precision Plus Protein™ Unstained Standard (Bio-Rad Laboratories, Inc., Hercules, CA, USA) was used as a

size marker for the molecular analysis of proteins. Analysis of protein bands and molecular weight (MW) estimates was performed using a Molecular Imager® Gel Doc™ XR+ System (Bio-Rad Laboratories, Inc., Hercules, CA, USA) and Quantity One® software (Bio-Rad Laboratories, Inc., Hercules, CA, USA). Accurate MW determinations and Peptide mass fingerprinting analysis was performed via mass spectrometry (MS) at the Mass Spectrometry Facility, Advanced Analysis Centre of the University of Guelph (Ontario, Canada).

Cell decoration with bacteriophage host recognition binding proteins. Confocal laser immunofluorescent microscopy was used to visualize the binding of the phage TG1 LTF protein Gp37 to *Y. enterocolitica* following methodology described by others (30). *Yersinia* strains grown in TSB at 25°C or 37°C for 24 hours were resuspended in wash buffer (50 mM Tris-HCl, pH 7.5) and 10 µl were spotted onto clean glass slides. After air-drying, the cells were fixed in a solution of 5 % gluteraldehyde for 10 min. and blocked with blocking buffer (5% BSA in 50 mM Tris-HCl buffer, pH 7.5) for 10 min. The slides were then incubated for 1 hour in a solution containing of 10 µg/mL of phage TG1 Gp37 (prepared in blocking buffer) followed by washing three times for 5 minutes in wash buffer. The slides were then incubated for 1 hour in anti-His_6 tag (HIS.H8) mouse monoclonal antibody solution prepared in blocking buffer (1:1000 dilution) and washed three times for 5 minutes with wash buffer. In a dark room, the slides were then incubated for 1 hour in goat anti-mouse IgG DyLight 488 polyclonal antibody solution (1:500) prepared in blocking buffer and washed three times for 5 minutes with wash buffer. The slides were air dried prior to analysis. Cells were imaged using an upright Leica DM 6000B confocal laser microscope connected to a Leica TCS SP5 system. Images were

collected digitally using Leica LAS AF Imaging Software and processed using ImageJ (31). To verify the specificity of the fluorescent signal, control samples were immunolabelled as above, with the omission of incubation with the primary antibody. All antibodies were acquired from Pierce Scientific, USA.

Genome sequences. The complete genome sequences of *Yersinia* phage vB_YenM_TG1 and vB_YenM_ϕR1-RT were deposited in the NCBI nucleotide database (GenBank) under the accession numbers KP202158 and HE956709, respectively. The RNA sequence data has been deposited to Gene Expression Omnibus (Acc. no GSE66516).

RESULTS

Phage Morphology. Phages ϕR1-RT and TG1 were negatively stained and examined by TEM. Both phages exhibit a prolate head with apparent icosahedral symmetry and a tubular contractile and rigid tail showing transverse striations (**Fig. 1**). The average dimension for the ϕR1-RT head is 82 ± 4 nm short edge-to-edge, 101 ± 5 nm vertex-to-vertex and the tail including the baseplate is on average 130 ± 7 nm long. The average dimension for the TG1 head is 91 ± 2 nm short edge-to-edge, 115 ± 6 nm vertex-to-vertex and the tail including the baseplate is on average 129 ± 1 nm long. Collectively, these morphological features indicate that these phages belong to the *Myoviridae* family.

Host specificity. The host range of phages TG1 and ϕR1-RT were determined by testing their lytic activity on 160 and 109 strains, respectively, belonging to thirteen *Yersinia* species, revealing virulence for *Y. enterocolitica* strains of serotypes O:1,

O:2, O:3, O:5, O:6, O:5,27, O:7,8, O:9 and some strains of serotype O:6,30 and O:6,31 while strains from other *Y. enterocolitica* serotypes and species within the genus *Yersinia* were resistant to phage infection (**Table 2**). TG1 and ϕR1-RT lysed their host when grown at 25°C but not at 37°C. Additionally, TG1 was unable to infect strains belonging to other 20 other genera (**Table S3**) demonstrating the phages' host range is restricted to *Y. enterocolitica.*

General features of the phage genomes. The genome of phage ϕR1-RT is 168,809 bp long with a GC content of 34.5%. The genome encodes 262 ORFs, of which 217 genes are encoded on the reverse strand (as displayed on the genetic map) and 45 genes on the forward strand (**Fig. S2**), with sizes ranging from 117 bp (product of 38 amino acids) to 3,738 bp (product of 1245 amino acids). The genome of TG1 is smaller than that of ϕR1-RT at 162,101 bp in length but with a similarly low GC content of 34.6%. TG1 also encodes 262 ORFs of which 223 genes are transcribed on the reverse strand (as displayed on the genetic map) and thirty-nine genes on the forward strand (**Fig. S3**), with sizes ranging from 114 bp (product of 37 amino acids) to 3,099 bp (product of 1032 amino acids). The GC content of these phages is significantly lower than that associated with the host with a GC content ranging from 47.1 ± 0.2 (32) to 48.5 ± 1.5 % (33). The genomes encode additionally four identical tRNA genes (Gly_{GGA}, Trp_{TGG}, Arg_{AGA}, Met_{ATG}) identified using tRNAScan (34) and ARAGORN (35).Constitutively low GC phage genomes are often supplemented with tRNA genes that, once expressed, enhance translation efficiency when infecting high GC content hosts (36). At the DNA level the TG1 genome shows 98% identity with a query coverage of 93%, for an overall DNA sequence identity of 91% with ϕR1-RT. All ORFs were screened using the BLASTP and PSI-BLAST algorithms (37, 38).

334 Based on protein homology, putative functions could be assigned to 121 (46%) gene
335 products of phage TG1 and 115 gene products (44 %) of phage ϕR1-RT. Most of the
336 identified homologs are conserved among T4-like phages and are either structural, or
337 involved in DNA replication, recombination, repair, or packaging. Thus, the phage T4
338 gene nomenclature was used to name these genes (**Table S4**).
339
340 **DNA replication, recombination, and repair.** Numerous genes were identified
341 within the phage ϕR1-RT and TG1 genomes that play a direct role in DNA
342 replication, recombination, and repair. Among the genes directly involved in DNA
343 replication are a DNA polymerase, a DNA ligase (Gp30), and three proteins with
344 helicase activity. The closest homologs to the phage TG1 and ϕR1-RT polymerases
345 are found in *Edwardsiella* phage PEi20 [BAQ22701.1] and *Enterobacteria* phage
346 RB69 [NP_861746.1]; all members of the *Myoviridae*. Among the helicases, Gp41
347 (or Dda) and UvsW homologs are involved in the reorganization of stalled DNA
348 replication forks (39). Other putative proteins identified include homologs to the
349 DNA polymerase sliding clamp loader complex Gp44/Gp62, sliding clamp accessory
350 protein Gp45, single-stranded DNA binding protein Gp32, DNA helicase loader
351 Gp59, and Gp61. In phage T4, the latter is a primase that interacts with helicase Gp41
352 to form a helicase-primase complex (or primosome). The primosome together with
353 the DNA helicase loader Gp59, unwinds the DNA template and primes DNA
354 synthesis on the discontinuous strand. Among the proteins involved in recombination
355 are type II topoisomerases Gp60 and Gp52, the recombination-related endonuclease
356 pair Gp46/Gp47, the Rec-A like recombination protein UvsX, and a single stranded
357 DNA binding protein, UvsY (40). Lastly, among the proteins involved in repair, a
358 DenV homolog and several RNA ligases were identified. DenV is an N-glycosylase

359 UV repair enzyme that excises pyrimidine dimers; the major UV-lesions of DNA,
360 while RNA ligases seal breaks in RNA and may also counteract host defense of
361 cleavage of specific tRNA molecules (41).
362
363 **Nucleotide metabolism.** Class I ribonucleotide reductases are responsible for the
364 inter-conversion of ribo- to deoxyribonucleotides and are represented by NrdA-B or
365 NrdE-F which require oxygen for activity, class II containing NrdJ, and the oxygen
366 sensitive class III represented by NrdG-H (42). In TG1 and ϕR1-RT, genes coding for
367 the aerobic ribonucleotide reductase complex subunits NrdA, NrdB, and NrdH were
368 identified. Additionally, NrdC genes were also located. Other genes identified that are
369 involved in nucleotide metabolism include: thymidylate synthase (*Td*), thymidine
370 kinase (*Tk*), dNMP kinase, dCMP deaminase (*Cd*), dihydrofolate reductase (*Fdr*),
371 dCTPase-dUTPase, and the exo-deoxyribonuclease *DexA* and endo-
372 deoxyribonuclease *DenA*. A combination of at least some of these genes is required to
373 supplement the intracellular pool of nucleotides for phage DNA and RNA synthesis
374 (41).
375
376 **Transcription.** Based on the genome maps presented (**Fig. S2 and S3**), phage TG1
377 and ϕR1-RT present a similar gene arrangement. A search for promoters based on
378 sequence similarity to the host consensus s70 promoter TTGACA(N15-18)TATAAT
379 with a 2 bp mismatch, identified 22 probable host promoters in the phage TG1
380 genome and 24 probable host promoters in the ϕR1-RT genome which probably
381 function in early transcription (**Tables S5 and S6**). Additionally, 15 of the putative
382 host promoters are located in the same relative genomic positions within each phage
383 genome. The genomic layout however, makes it clear that there must be additional

promoters functioning to direct the transition from host to viral metabolism. A search for phage-specific promoters using PHIRE (43) and by analysis of sequences of 100 bp in length upstream of each ORF and submitting them to MEME (44), did not yield additional promoters that could be annotated with confidence. A search for putative rho-independent transcription terminators using ARNold (45, 46) yielded 21 putative terminators in the phage TG1 genome (**Table S7**) and 24 in the phage ϕR1-RT genome (**Table S8**). Nevertheless, the presence phage T4 homologs involved in the transcription of late genes: RegA, Gp33, and the sigma factor for late transcription Gp55, suggest that the mechanism for controlling late transcription is similarly complex (41). Likewise, the presence of repressor and translational regulatory protein homologs involved in middle and late transcription including: RegB, DsbA, Alc, MotA, and AsiA, lend further support to this suggestion.

Morphogenesis. The putative structural proteins of TG1 and ϕR1-RT are homologous to existing phage proteins of the T4 supergroup of viruses (**Table S4**). Among the putative phage structural genes, the phage head is likely composed of the major capsid protein Gp23 and the phage capsid vertex protein Gp24. The prohead precursor and scaffolding proteins Gp68, and Gp67 as well as internal head proteins ipIII and ipII were also identified. Lastly the head portal vertex protein Gp20 that is connected to the neck and through which DNA enters during packaging and exits during infection was also identified. The whiskers and neck are composed of fibritin (*wac*) and the head completion proteins Gp13 and Gp14. The tail proteins include the tail sheath terminator Gp3, the tail completion protein Gp15, the tail sheath subunit Gp18, and the tail tube subunit Gp19. Proteins that form the baseplate wedge subunits and tail pins that then go on to associate with the central hub to form the viral

409 baseplate include: Gp5, Gp6, Gp7, Gp8, Gp9, Gp10, Gp11, Gp25, Gp27, Gp28, and
410 Gp53. Among these, Gp5 (*ORF150*) contains a predicted bacteriophage T4-like
411 lysozyme domain (cd00735) or Phage lysozyme domain (pfam00959), which aids
412 penetration through the peptidoglycan layer during the initial infection process. In
413 phage T4, Gp8 and Gp9 connect the long tail fibers of the virus to the baseplate and
414 trigger tail contraction after viral attachment to a host cell, while Gp11 connects the
415 short tail fiber protein Gp12 (*ORF159*) to the baseplate (47). The baseplate wedge
416 subunit Gp25, forms a structural component of the outer wedge of the baseplate that
417 has lysozyme activity, evident by the presence of conserved Gene 25-like lysozyme
418 domain (pfam04965). Based on homology and gene synteny the proteins forming the
419 long tail fibers in TG1 and ϕR1-RT are composed of the tail fiber proximal subunit
420 Gp34, the tail fiber connector or hinge protein Gp35, the proximal tail fiber protein
421 Gp36, and the distal tail fiber protein Gp37 (47). A variety of chaperones or assembly
422 catalysts involved in morphogenesis were also discovered. Head formation
423 chaperones include the capsid vertex assembly chaperone, the prohead assembly
424 proteins Gp21 and Gp22, as well as the head assembly chaperone protein Gp31.
425 Chaperones involved in tail formation include the baseplate hub assembly proteins
426 Gp26 and Gp51. Chaperones for tail fiber assembly include gp57A, gp57B, and
427 Gp38.
428
429 **Host cell recognition elements.** In phage T4, phage tail associated receptor-binding
430 proteins (RBPs) Gp37 and Gp12 are necessary for host cell recognition, attachment,
431 and initiation of infection. In the phage TG1 and ϕR1-RT genomes *ORF250* codes for
432 a putative RBP protein of 609 amino acid residues and 503 amino acid residues in
433 length, respectively, sharing 60% overall sequence identity. These proteins also share

40% sequence identity to the distal long tail fiber RBP of *Cronobacter* phage vB_CsaM_GAP161 [YP_006986537.1] and are homologs to the long tail fiber RBP Gp37 of phage T4 [AJC64544.1]. An alignment of these two proteins reveals a high degree of conservation at the N-terminus associated with the proximal tail fiber, as well as at the C-terminus associated with host recognition (**Fig. S4**). More specifically, the C-terminal 63 amino acids present a 95% sequence identity. Similarly, in the phage TG1 and ϕR1-RT genomes, *ORF159* codes for the short tail fiber (STF) protein Gp12, both of 446 amino acid residues in length and which are almost identical to each other, sharing 95% overall sequence identity (**Fig. S5**). These proteins are homologous to the STF protein Gp12 of phage T4 [NP_049770.1].

DNA Packaging. In phage TG1, *ORF164* and *ORF165/ORF167* genes code for the small (TerS) and large (TerL) DNA packaging subunits respectively of a phage terminase protein complex (or holoterminase) that initiates, drives, and terminates translocation of phage DNA into proheads (48). The homologous genes in phage ϕR1-RT are represented by *ORF165* (TerS) and *ORF166/ORF168* (TerL). Usually, *terS* and *terL* are arranged side by side but in phage TG1 and ϕR1-RT two ORFs homologous with *terL* are found. *ORF165* in TG1 and *ORF166* in ϕR1-RT show sequence similarity to the N-terminus of the phage T4 TerL. Likewise *ORF167* in TG1 and *ORF168* in ϕR1-RT show sequence similarity to the C-terminus of phage T4 TertL. BLASTX analysis (37, 38) reveals the *terL* gene in both phages is interrupted by a transposase (PHA02552). Additionally, during packaging, the DNA ends are also protected against host RecBCD nuclease action by Gp2, the DNA end protector protein (49); identified in phage TG1 and in ϕR1-RT as the product of *ORF146*.

458 **Homing Endonucleases.** Homing endonuclease genes (HEGs) are not genuine phage
459 DNA, but rather belong to intron associated selfish DNA elements (50) and are
460 commonly found interspersed throughout *Myoviridae* genomes (41). Among the
461 HEGs identified in phage TG1, *ORF148* and *ORF232* exhibit similarity to shortened
462 helix-turn-helix (HN-H) endonucleases, and *ORF9*, *ORF43*, and *ORF66* to GIY-YIG
463 group I intron endonucleases. BLASTX analysis (37, 38) reveals Gp47
464 (recombination-related endonuclease II) is divided by *ORF66* which contains the
465 HEG. Likewise, the gene coding for UvsX is intersected by *ORF43*, which contains
466 the HEG. *ORF232* also divides the *NrdA* gene. In phage ϕR1-RT five HEGs are also
467 found throughout the genome of which only *ORF148* is homologous to the helix-turn-
468 helix (HN-H) endonuclease that is also located in phage TG1 between Gp4 and Gp53.
469 *ORF20*, *ORF51*, *ORF163*, and *ORF234* exhibit similarity to GIY-YIG group I intron
470 endonucleases, none of which interrupt or intersect other ϕR1-RT genes.
471
472 **Lysis.** The final stage of the phage lytic cycle involves the degradation of the bacterial
473 cell wall and release of progeny phages induced by the effect a pore producing
474 protein, the holin, and a peptidoglycan degrading enzyme, the endolysin (51). In TG1,
475 *ORF127* and *ORF122* in ϕR1-RT each encode an obvious endolysin containing a
476 bacteriophage T4-like lysozyme protein domain (pfam00959) and phage-related
477 muramidase (COG3772). Access of the endolysin to the cell wall occurs through the
478 presence of the holin. Holins are small phage encoded proteins characterized by the
479 presence of TMDs, accumulating in the cytoplasmic membrane during infection until
480 suddenly at a specific time, trigger to form lethal lesions resulting in destruction of the
481 cell wall (51, 52). A search for the TG1 and ϕR1-RT holins revealed the putative
482 product of their respective *ORF253* gene contains a predicted t-holin domain

(pfam11031) with 70% identity to the phage holin of *Enterobacteria* phage CC31 [YP_004010117.1]. The protein sequences are predicted to contain a single TMD spanning aa interval 30-49, as well as a large C-terminal periplasmic domain spanning the aa residue from position 50 to the end terminal amino acid at position 218; a characteristic bitopic topology found in the holin proteins of T4-like phages (53) Moreover, as in phage T4 the putative holin gene is separated from the endolysin gene. An additional search for *Rz/Rz1* genes coding for transmembrane spanins involved in the disruption of the outer membrane of the host was also conducted based on gene arrangement and membrane localization signals (54). The search revealed two candidate genes, *ORF225* and *ORF224* in phage TG1 and *ORF227* and *ORF226* in phage ϕR1-RT, homologous to phage T4 pseT.3 (*Rz*) and pseT.2 (*Rz1*), respectively. As in phage T4, the *Rz/Rz1* genes are adjacent to each other, arranged with overlapping stop and start codons, and additionally no part of the *Rz1* sequence is embedded within the *Rz* coding region (54). In TG1 and ϕR1-RT, *Rz* possesses a single amino-terminal TMD, and *Rz1* encodes an outer membrane lipoprotein based on the presence of a signal peptidase II (SPII) cleavage site located between amino acid residues 16 and 17 as predicted by LipoP (54, 55). Lastly, the presence of phage T4 homologs to *rI* lysis inhibition regulator membrane protein and *rIII* lysis inhibitor accessory protein in TG1 and ϕR1-RT suggest the potential for lysis inhibition (LIN) following superinfection (56, 57).

Phylogeny of TG1. It is interesting to note that very similar bacteriophages with an overall DNA sequence identity of 91% were isolated from such different locations and sources, as phages TG1 and ϕR1-RT were isolated in Canada from pig manure, and in Finland from raw sewage, respectively. Moreover, less than 34% overall DNA

508 similarity exists with their closest neighbours within the *Myoviridae* (**Table S9**). The
509 relatedness of these two phages was further explored using progressiveMauve (**Fig. 2**)
510 (58, 59); CoreGenes (60, 61) which the Bacterial and Archaeal Virus Subcommittee
511 of the International Committee on Taxonomy of Viruses (ICTV) has extensively used
512 to compare the proteomes of viruses; and by phylogenetic analysis of their whole
513 genome sequences (**Fig. S6**) and their large terminase subunit protein sequences (**Fig.**
514 **3**). It is evident from phylogenetic analyses that TG1 and ϕR1-RT form a distinct
515 taxonomic clade among their closest neighbours. Based on these observations and
516 using a 95% DNA sequence identity as the criterion for demarcation for a species, a
517 new genus named *Tg1virus* with phages TG1 and ϕR1-RT as member species was
518 proposed to the ICTV (approved in 2016 and pending ratification).
519
520 **Growth curves.** In order to study the efficiency of phage infection at different
521 temperatures bacterial growth after phage infection with ϕR1-RT was measured. Host
522 bacterial strain was grown at selected temperatures with addition of different phage
523 stock dilutions. Bacterial growth was followed for 3 d at 4°C, 2d at 10°C and 16°C
524 and 1 d at 22°C and 37°C. Lysis of the bacterial cultures was observed at 4°C, 10°C,
525 16°C and 22°C, whereas at 37°C the bacteria were not significantly affected even
526 with the highest initial phage concentrations (**Fig. 4**, panels A). The onset time of
527 lysis depended on the temperature and initial phage titer. At 4°C the bacterial culture
528 started to lyse after 56-60 h, at 10°C already at 16 h with highest phage titer and at
529 24-28 h with the lowest phage titer. The corresponding times for 16°C and 22°C were
530 6 and 12 h. While the lysis at 10°C and 16°C was complete, at 22°C strong regrowth
531 after the initial lysis took place. At 4°C the 3 d incubation time was not long enough
532 to follow the lysis to completion. Under all tested conditions negative (medium only)

controls showed no increase in the absorbance, whereas the positive (bacteria only) controls presented the normal bacterial growth pattern.

Transduction. To study the transducing potential of ϕR1-RT, transduction of the Km^R and urease-negative phenotype of strain YeO3-*hfq*::Km to the Km^S and urease-positive wild type strain 64741/76 was assayed. Repression of urease activity is one of the phenotypes of the *hfq* mutant (62) and could be used to confirm the transduction of the *hfq::*Km allele. The transduction assays were performed in 10 parallel tubes using a MOI of 0.006. A total of 6.6×10^8 PFU from the transducing lysate resulted in a total of 3 Km^R urease-negative colonies that were confirmed by PCR. From this the calculated transduction frequency in the experiment was 4.5×10^{-7} transductants per PFU.

Identification of the phage receptors – pseudolysogeny. As the LPS and protein profiles of the phage resistant mutants YeO3-ϕR1-RT-2, YeO3-ϕR1-RT-7, and YeO3-ϕR1-RT-9 did not differ from those of the wild type bacteria (data not shown) the genomic DNA of the mutant and the wild type strains were sequenced. The *de novo* assembly results showed that the total scaffold sizes of the assembled genomes of the three mutants were ~165-173 kb larger than that of the wild type parental strain (**Table S10**). This suggested that the mutants carried extra DNA and the size matched very close to the size of phage ϕR1-RT genome (168,809 bp). This immediately raised the possibility that the phage had lysogenized the host and would reside as a prophage. In all three draft genomes the phage genome sequence was indeed identified and in all it formed the scaffold 4.1 with almost identical sizes (**Table S10**). Significantly, in all three cases the scaffold sequences were 100% identical to phage

558 ϕR1-RT sequence without any flanking host sequences, suggesting that the phage
559 genome resided in these bacteria as an autonomous replicating unit in a state known
560 as pseudolysogeny. Such state has been described for T4-like phages (63).
561
562 **Identification of the phage receptors – transposon insertion library screening.** As
563 selection of spontaneous phage resistant mutants seemed to favor pseudolysogeny we
564 decided to use a different approach. A *CatMu*-transposon library of strain YeO3-R1
565 (26) was exposed the ϕR1-RT for 2 hr and the surviving phage-resistant mutants were
566 grown on LA-Clm plates. The recovered colonies were tested to be true ϕR1-RT
567 resistant mutants. In order to exclude pseudolysogens, the clones were screened with
568 ϕR1-RT specific PCR, and negative ones were further analysed by *CatMu*-specific
569 arbitrary PCR to identify the CatMu insertion site (26). For four of the candidates the
570 transposon insertion site was identified as gene *Y11_04441* of the *Y. enterocolitica*
571 O:3 strain Y11 genome (NC_017564.1). In strain Y11 genome the gene was
572 annotated to encode for the outer membrane porin OmpC, however, in all other *Y.*
573 *enterocolitica* genomic sequences as OmpF, therefore we opted to use OmpF. To
574 confirm that OmpF is the ϕR1-RT receptor, one of the mutants YeO3-R1-Cat17 was
575 complemented with the wild type *ompF* gene either *in trans* with plasmid
576 pTM100_OmpF or *in cis* by suicide plasmid pSW25T_OmpF. Both of these
577 approaches resulted in regaining the phage sensitivity thus confirming that OmpF
578 serves as ϕR1-RT receptor.
579
580 **The LPS inner core heptose region functions as a receptor.** Adsorption
581 experiments were carried out to study the ability of ϕR1-RT and TG1 to interact with
582 *Y. enterocolitica* O:3 derivatives differing mainly in their LPS composition (**Fig. 5**). A

583 short 5 min adsorption time was used as it produced highest resolution between the
584 strains. A general observation was that TG1 adsorbed faster than ϕR1-RT. Both
585 phages showed negligible adsorption to YeO3-c-R1-Cat17, the *ompF* mutant strain
586 and adsorbed well to both *ompF*-complemented strains. Both phages showed reduced
587 but clear adsorption to the pseudolysogen, indicating changes in abundance or
588 exposure of the phage receptor(s). Finally, the adsorption to the inner core mutants
589 decreased with the truncation of the core oligosaccharide suggesting that the inner
590 core heptoses are part of the secondary receptor (**Fig. 5**).
591
592 **Temperature-dependence of *ompF* expression.** We then wondered whether the
593 temperature-dependent sensitivity of *Y. enterocolitica* O:3 could be due to *ompF*
594 regulation. The expression of *ompF* under different growth temperatures was analysed
595 from RNA-sequencing and quantitative proteomics (LC-MS/MS) data. The
596 transcriptomic data showed an inverse correlation between the expression of *ompF*
597 and the temperature of incubation (**Fig. 4**, panels B). Consistently, the quantitative
598 proteomics demonstrated much higher abundance of the OmpF protein in the 22°C
599 sample when compared to the 37°C sample, where the abundance barely exceeded the
600 threshold of identification (**Fig. 4**, panels B).
601
602 **In vitro expression of the long tail fiber protein Gp37 of phage TG1.** In this study,
603 co-expression with the phage encoded chaperones Gp38 (required for
604 oligomerization) and Gp57A, which is also thought to participate in assembly (64, 65)
605 was utilized in an attempt to synthesize the native form of distal long tail fiber protein
606 of phage TG1 as previously described for the production of Gp37 from phage T4 (28).
607 SDS-PAGE demonstrated that an oligomer of approximately 210 kDa was obtained

608 when Gp37 was co-expressed with Gp38 in a bicistronic plasmid (pCDF-Duet-1
609 Gp37-Gp38) or when this same plasmid was co-expressed with Gp57A (**Fig. S7**,
610 lanes 3 and 5). Under reduced conditions Gp37 appears as a monomer of
611 approximately 70 kDa in size (**Fig. S7**, lanes 4 and 6). This estimate is close to the
612 predicted molecular mass of the recombinant phage TG1 Gp37 determined via MS at
613 approximately 68.050 kDa. Peptide mass fingerprinting confirmed the identity of the
614 protein (**Fig. S8**). Based on the protein expression results obtained, it appears that in
615 phage TG1 only the Gp38 chaperone is essential and the general chaperone Gp57A is
616 not required for *in-vitro* protein folding of Gp37 as has been reported for phage T4
617 (28). The formation of higher molecular weight oligomers of phage TG1 Gp37 is
618 consistent with previous reports that describe RBPs of phages present as homotrimers
619 in solution migrating in the SDS-PAGE with a mobility that corresponds to that of
620 oligomeric forms (28, 66–68).
621
622 **Confirmation of host binding specificity**. Host binding specificity was then tested
623 through immunolabeling of bacterial cells with phage TG1 LTF protein Gp37
624 followed by detection with anti-His_6 antibodies and DyLight 488 conjugated
625 secondary antibodies. Consistent with the temperature dependent infection of phage
626 TG1, the application of the LTF protein Gp37 to *Y. enterocolitica* cells showed
627 decoration of the surface of *Y. enterocolitica* O:3, O:5,27, and O:9 cells when these
628 were grown at 25°C but not at 37°C (**Fig. 6**). Notably, binding was more apparent
629 near the apex of the cells which is also reported to occur in other phages such as λ,
630 T4, T7, KVP40 and ϕA1122, preferentially infecting cells at the poles (69).
631
632 **DISCUSSION**

633 Among bacteriophages, the C-terminus of RBPs involved in ligand interactions
634 usually exhibits considerable sequence divergence, thus providing diversity in host
635 specificity. In the case of ϕR1-RT and TG1, the high sequence identity at the C-
636 terminus of their long tail fiber and short tail fiber proteins may account for the
637 striking similarity in virulence of these two phages for *Y. enterocolitica*. Notably,
638 phage ϕR1-RT shows virulence to strains of the same serotypes as phage TG1. Based
639 on adsorption experiments, the outer membrane protein OmpF and the inner core
640 heptosyl residues of the LPS serve as phage receptors for phage TG1 and ϕR1-RT. It
641 is worth noting however, that the *E. coli* strain DH10B/pTM100_OmpF was not
642 sensitive to ϕR1-RT. We reasoned that this could be due to poor expression of *Y.*
643 *enterocolitica* OmpF in *E. coli* or more likely that the LPS inner core, known to be
644 used by T4-like phages as the secondary receptor (76, 77) was not compatible. The
645 inner core structures of *E. coli* and *Y. enterocolitica* differ substantially potentially
646 explaining this result.
647
648 Multiple lines of evidence suggest OmpF is the primary host range determinant for
649 these two bacteriophages. First, a multiple alignment of OmpF amino acid sequences
650 of *Y. enterocolitica* (from a BLASTP search of sequence databases using the O:3
651 OmpF sequence as query) suggest the restricted host range of these phages among *Y.*
652 *enterocolitica* serotypes could be due to OmpF. The alignment provided a distribution
653 of conserved amino acid residues and the presence of regions with high and low
654 homologies, which coincide with eight transmembrane domains and eight "external"
655 loops, respectively of the topology of the OmpF porin from *E. coli* (70, 71). The
656 search and alignment of the sequences (**Fig. S9**) revealed that the OmpF sequences of
657 the ϕR1-RT sensitive serotypes are 100% identical. The most dramatic differences

658 between the serotypes map to loop 4. In the alignment most close to the O:3 sequence
659 is the serotype O:7,8,19 OmpF that is 96% identical to O:3 and may still be sensitive
660 to ϕR1-RT; in it the loop 4 sequence differs least, while in others differences are
661 bigger and also accumulate in other loops, mainly in loops 5, 6, and 7 (**Fig. S9**). The
662 porin loops are plausible binding sites for bacteriophages as demonstrated by the
663 interaction of *E. coli* OmpF and K20 phages which bind to the L5, L6, and L7
664 external loops (72–74). Thus it is likely that the loop 4 sequence is targeted by the
665 ϕR1-RT or TG1 receptor binding proteins, however, experimental evidence is
666 necessary to confirm this. Secondly, RNA-sequencing and quantitative proteomics
667 data, the analysis of growth curves of *Y. enterocolitica* infected with ϕR1-RT at
668 various temperatures (4°C to 37°C), as well as phage host range analysis results
669 conducted at 25°C and 37°C clearly indicate that the failure of ϕR1-RT and TG1 to
670 infect *Y. enterocolitica* O:3 at 37°C is due to the strong repression of the *ompF* gene.
671 The temperature dependent expression of OmpF results also agree with a previous
672 study, where two-dimensional gel electrophoresis of whole-cell proteins of *Y.*
673 *enterocolitica* cultured at 25°C and 37°C suggested that OmpF is downregulated
674 when the bacteria were cultured at 37°C (75). Consistent with this observation, the
675 application of immunolabelled phage TG1 receptor binding protein Gp37 to *Y.*
676 *enterocolitica* cells showed decoration of the surface of *Y. enterocolitica* O:3, O:5,27,
677 and O:9 cells when these were cultured at 25°C but not at 37°C. The decoration of the
678 cell surface agrees with a high level expression of this major outer membrane protein
679 class depending on the bacterial species and the environmental conditions, which can
680 reach about 10^4–10^6 copies per cell (72). It is reasonable to suggest then that the
681 phage TG1 distal long tail fiber protein Gp37 (and by extension, its homolog in ϕR1-
682 RT) is specifically involved in binding to OmpF while presumably, the short tail fiber

683 protein Gp12 binds to the inner core of LPS, as is reported to occur in other T even
684 phages as a secondary receptor (76, 77).
685
686 The *in vitro* temperature dependent infection of these two highly related phages,
687 questions their potential use as biocontrol or therapeutic agents as has been suggested
688 for the temperate *Yersinia* phage PY100 (18, 78). On the other hand, it is not known
689 whether the *ompF* gene is expressed in vivo justifying further studies towards finding
690 that out. However, due to their marked specificity for the epidemiological relevant *Y.*
691 *enterocolitica* serotypes O:3, O:5,27, and O:9, these phages may prove useful for
692 diagnostic purposes. In addition, the successful synthesis of the long tail fiber of
693 phage TG1 opens up the possibility of its use as a probe as well as for the production
694 of suitable amounts of protein for X-ray crystallography to elucidate its atomic
695 structure or co-crystallization with its receptor OmpF to shed light on specific host
696 cell receptor-virus interactions.
697
698 **ACKNOWLEDGEMENTS**
699 Karolina Grabowska and Sofia Itkonen are thanked for help in screening the CatMu
700 transposon library.
701
702 **FUNDING INFORMATION**
703 This research was supported by the Ontario Ministry of Agriculture (OMAF) Food
704 Safety Research Program (research grant SF6075 to JAO) and the Academy of
705 Finland (grants 114075 and 288701 to MS). The funders had no role in study design,
706 data collection and interpretation, or the decision to submit the work for publication.

REFERENCES

1. **Bottone EJ**. 1999. *Yersinia enterocolitica*: overview and epidemiologic correlates. Microbes Infect **1**:323–333.

2. **Fredriksson-Ahomaa M**, **Stolle A**, **Korkeala H**. 2006. Molecular epidemiology of *Yersinia enterocolitica* infections. FEMS Immunol Med Microbiol **47**:315–329.

3. **Fukushima H**, **Shimizu S**, **Inatsu Y**. 2011. *Yersinia enterocolitica* and *Yersinia pseudotuberculosis* Detection in Foods. J Pathog **2011**:735308.

4. **McNally A**, **Thomson NR**, **Reuter S**, **Wren BW**. 2016. "Add, stir and reduce": *Yersinia* spp. as model bacteria for pathogen evolution. Nat Rev Microbiol **14**:177–190.

5. **Verhaegen J**, **Charlier J**, **Lemmens P**, **Delmee M**, **Van Noyen R**, **Verbist L**, **Wauters G**. 1998. Surveillance of human *Yersinia enterocolitica* infections in Belgium: 1967-1996. Clin Infect Dis **27**:59–64.

6. **Wren BW**. 2003. The yersiniae-a model genus to study the rapid evolution of bacterial pathogens. Nat Rev **1**:55–64.

7. **Wauters G**, **Kandolo K**, **Janssens M**. 1987. Revised biogrouping scheme of *Yersinia enterocolitica*. Contrib Microbiol Immunol **9**:14–21.

8. **Aussel L**, **Therisod H**, **Karibian D**, **Perry MB**, **Bruneteau M**, **Caroff M**. 2000. Novel variation of lipid A structures in strains of different *Yersinia* species. FEBS Lett **465**:87–92.

9. **Bruneteau M**, **Minka S**. 2003. Lipopolysaccharides of bacterial pathogens from the genus *Yersinia*: a mini-review. Biochimie **85**:145–152.

10. **ECDC**. 2015. Surveillance of seven priority food- and waterborne diseases in the EU/EEA. Stockholm, Sweden.

11. **Kiljunen S**, **Vilen H**, **Pajunen M**, **Savilahti H**, **Skurnik M**. 2005. Nonessential genes of phage phiYeO3-12 include genes involved in adaptation to growth on *Yersinia enterocolitica* serotype O:3. J Bacteriol **187**:1405–1414.

12. **Pajunen M**, **Kiljunen S**, **Skurnik M**. 2000. Bacteriophage phiYeO3-12, specific for *Yersinia enterocolitica* serotype O:3, is related to coliphages T3 and T7. J Bacteriol **182**:5114–5120.

13. **Pajunen MI**, **Kiljunen SJ**, **Söderholm ME**, **Skurnik M**, **Soderholm ME**, **Skurnik M**. 2001. Complete genomic sequence of the lytic bacteriophage

phiYeO3-12 of *Yersinia enterocolitica* serotype O:3. J Bacteriol **183**:1928–1937.

14. **Leon-Velarde CG**, **Kropinski AM**, **Chen S**, **Abbasifar A**, **Griffiths MW**, **Odumeru JA**. 2014. Complete genome sequence of bacteriophage vB_YenP_AP5 which infects *Yersinia enterocolitica* of serotype O:3. Virol J **11**:188.

15. **Hertwig S**, **Klein I**, **Schmidt V**, **Beck S**, **Hammerl JA**, **Appel B**. 2003. Sequence analysis of the genome of the temperate *Yersinia enterocolitica* phage PY54. J Mol Biol **331**:605–622.

16. **Kiljunen S**, **Hakala K**, **Pinta E**, **Huttunen S**, **Pluta P**, **Gador A**, **Lonnberg H**, **Skurnik M**. 2005. Yersiniophage phiR1-37 is a tailed bacteriophage having a 270 kb DNA genome with thymidine replaced by deoxyuridine. Microbiology **151**:4093–4102.

17. **Skurnik M**, **Hyytiäinen HJ**, **Happonen LJ**, **Kiljunen S**, **Datta N**, **Mattinen L**, **Williamson K**, **Kristo P**, **Szeliga M**, **Kalin-Mänttäri L**, **Ahola-Iivarinen E**, **Kalkkinen N**, **Butcher SJ**. 2012. Characterization of the genome, proteome, and structure of yersiniophage ϕR1-37. J Virol **86**:12625–42.

18. **Schwudke D**, **Ergin A**, **Michael K**, **Volkmar S**, **Appel B**, **Knabner D**, **Konietzny A**, **Strauch E**. 2008. Broad-host-range Yersinia phage PY100: genome sequence, proteome analysis of virions, and DNA packaging strategy. J Bacteriol **190**:332–342.

19. **Skurnik M**. 1999. Genetics of Bacterial Polysaccharides, p. 23–51.. *In* Goldberg, J (ed.), . CRC Press, Boca Raton, FL.

20. **Al-Hendy A**, **Toivanen P**, **Skurnik M**. 1991. Expression cloning of *Yersinia enterocolitica* O:3 rfb gene cluster in Escherichia coli K12. Microb Pathog **10**:47–59.

21. **Baker PM**, **Farmer JJ**. 1982. New bacteriophage typing system for *Yersinia enterocolitica*, *Yersinia kristensenii*, *Yersinia frederiksenii*, and *Yersinia intermedia*: correlation with serotyping, biotyping, and antibiotic susceptibility. J Clin Microbiol **15**:491–502.

22. **Al-Hendy A**, **Toivanen P**, **Skurnik M**. 1992. Lipopolysaccharide O side chain of *Yersinia enterocolitica* O:3 is an essential virulence factor in an orally infected murine model. Infect Immun **60**:870–875.

23. **Skurnik M**. 1984. Lack of correlation between the presence of plasmids and fimbriae in *Yersinia enterocolitica* and *Yersinia pseudotuberculosis*. J Appl Bacteriol **56**:355–363.

24. **Kutter E**. 2009. Phage Host Range and Efficiency of Plating, p. 141. *In* Clokie, M, Kropinski, AM (eds.). Humana Press.

25. **Biedzka-Sarek M**, **Venho R**, **Skurnik M**. 2005. Role of YadA, Ail, and

Lipopolysaccharide in Serum Resistance of *Yersinia enterocolitica* Serotype O:3. Infect Immun **73**:2232–2244.

26. **Pinta E**, **Li Z**, **Batzilla J**, **Pajunen M**, **Kasanen T**, **Rabsztyn K**, **Rakin A**, **Skurnik M**. 2012. Identification of three oligo-polysaccharide-specific ligases in *Yersinia enterocolitica.* Mol Microbiol **83**:125–36.

27. **Pajunen M**, **Kiljunen S**, **Skurnik M**. 2012. Construction and Screening of a Transposon Insertion Library of *Yersinia enterocolitica* (YeO3-R1). Bio-protocol **2**.

28. **Bartual SG**, **Garcia-Doval C**, **Alonso J**, **Schoehn G**, **van Raaij MJ**. 2010. Two-chaperone assisted soluble expression and purification of the bacteriophage T4 long tail fibre protein Gp37. Protein Expr Purif **70**:116–121.

29. **Laemmli UK**. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**:680–5.

30. **Javed MA**, **Poshtiban S**, **Arutyunov D**, **Evoy S**, **Szymanski CM**. 2013. Bacteriophage receptor binding protein based assays for the simultaneous detection of *Campylobacter jejuni* and *Campylobacter coli*. PLoS One **8**:e69770.

31. **Schneider CA**, **Rasband WS**, **Eliceiri KW**. 2012. NIH Image to ImageJ: 25 years of image analysis. Nat Methods **9**:671–5.

32. **Daligault HE**, **Davenport KW**, **Minogue TD**, **Bishop-Lilly KA**, **Broomall SM**, **Bruce DC**, **Chain PS**, **Coyne SR**, **Frey KG**, **Gibbons HS**, **Jaissle J**, **Koroleva GI**, **Ladner JT**, **Lo C-C**, **Munk C**, **Palacios GF**, **Redden CL**, **Rosenzweig CN**, **Scholz MB**, **Johnson SL**. 2014. Whole-Genome *Yersinia* sp. Assemblies from 10 Diverse Strains. Genome Announc **2**.

33. **Brenner DJ**, **Ursing J**, **Bercovier H**, **Steigerwalt AG**, **Fanning GR**, **Alonso JM**, **Mollaret HH**. 1980. Deoxyribonucleic acid relatedness in *Yersinia enterocolitica* and *Yersinia enterocolitica*-like organisms. Curr Microbiol **4**.

34. **Schattner P**, **Brooks AN**, **Lowe TM**. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. Nucleic Acids Res **33**:W686–9.

35. **Laslett D**, **Canback B**. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res **32**.

36. **Miller ES**, **Kutter E**, **Mosig G**, **Arisaka F**, **Kunisawa T**, **Rüger W**. 2003. Bacteriophage T4 genome. Microbiol Mol Biol Rev **67**:86–156.

37. **Altschul SF**, **Gish W**, **Miller W**, **Myers EW**, **Lipman DJ**. 1990. Basic local alignment search tool. J Mol Biol **215**:403–10.

38. **Altschul SF**, **Madden TL**, **Schäffer AA**, **Zhang J**, **Zhang Z**, **Miller W**,

Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25**:3389–402.

39. **Manosas M**, **Perumal SK**, **Croquette V**, **Benkovic SJ**. 2012. Direct observation of stalled fork restart via fork regression in the T4 replication system. Science **338**:1217–20.

40. **Liu J**, **Morrical SW**. 2010. Assembly and dynamics of the bacteriophage T4 homologous recombination machinery. Virol J **7**:357.

41. **Petrov VM**, **Ratnayaka S**, **Nolan JM**, **Miller ES**, **Karam JD**. 2010. Genomes of the T4-related bacteriophages as windows on microbial genome evolution. Virol J **7**:292.

42. **Lundin D**, **Torrents E**, **Poole AM**, **Sjöberg B-M**. 2009. RNRdb, a curated database of the universal enzyme family ribonucleotide reductase, reveals a high level of misannotation in sequences deposited to Genbank. BMC Genomics **10**:589.

43. **Lavigne R**, **Sun WD**, **Volckaert G**. 2004. PHIRE, a deterministic approach to reveal regulatory elements in bacteriophage genomes. Bioinformatics **20**.

44. **Bailey TL**, **Boden M**, **Buske FA**, **Frith M**, **Grant CE**, **Clementi L**, **Ren J**, **Li WW**, **Noble WS**. 2009. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res **37**:W202–8.

45. **Gautheret D**, **Lambert A**. 2001. Direct RNA motif definition and identification from multiple sequence alignments using secondary structure profiles. J Mol Biol **313**:1003–11.

46. **Macke TJ**, **Ecker DJ**, **Gutell RR**, **Gautheret D**, **Case DA**, **Sampath R**. 2001. RNAMotif, an RNA secondary structure definition and search algorithm. Nucleic Acids Res **29**:4724–35.

47. **Leiman PG**, **Arisaka F**, **van Raaij MJ**, **Kostyuchenko VA**, **Aksyuk AA**, **Kanamaru S**, **Rossmann MG**. 2010. Morphogenesis of the T4 tail and tail fibers. Virol J **7**:355.

48. **Mitchell MS**, **Rao VB**. 2006. Functional analysis of the bacteriophage T4 DNA-packaging ATPase motor. J Biol Chem **281**:518–27.

49. **Lipinska B**, **Rao AS**, **Bolten BM**, **Balakrishnan R**, **Goldberg EB**. 1989. Cloning and identification of bacteriophage T4 gene *2* product Gp2 and action of Gp2 on infecting DNA in vivo. J Bacteriol **171**:488–97.

50. **Bell-Pedersen D**, **Quirk S**, **Clyman J**, **Belfort M**. 1990. Intron mobility in phage T4 is dependent upon a distinctive class of endonucleases and independent of DNA sequences encoding the intron core: mechanistic and evolutionary implications. Nucleic Acids Res **18**:3763–70.

51. **Wang IN**, **Smith DL**, **Young R**. 2000. Holins: the protein clocks of bacteriophage infections. Annu Rev Microbiol **54**:799–825.

52. **Dewey JS**, **Savva CG**, **White RL**, **Vitha S**, **Holzenburg A**, **Young R**. 2010. Micron-scale holes terminate the phage infection cycle. Proc Natl Acad Sci U S A **107**:2219–23.

53. **Tran TAT**, **Struck DK**, **Young R**. 2005. Periplasmic domains define holin-antiholin interactions in T4 lysis inhibition. J Bacteriol **187**:6631–40.

54. **Summer EJ**, **Berry J**, **Tran TA**, **Niu L**, **Struck DK**, **Young R**. 2007. Rz/Rz1 Lysis Gene Equivalents in Phages of Gram Negative Hosts. J Mol Biol **373**.

55. **Juncker AS**, **Willenbrock H**, **Von Heijne G**, **Brunak S**, **Nielsen H**, **Krogh A**. 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci **12**:1652–62.

56. **Doermann AH**. 1948. Lysis and Lysis Inhibition with *Escherichia coli* Bacteriophage. J Bacteriol **55**:257–76.

57. **Burch LH**, **Zhang L**, **Chao FG**, **Xu H**, **Drake JW**. 2011. The bacteriophage T4 rapid-lysis genes and their mutational proclivities. J Bacteriol **193**:3537–45.

58. **Darling AC**, **Mau B**, **Blattner FR**, **Perna NT**. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome **14**.

59. **Darling AE**, **Mau B**, **Perna NT**. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One **5**:e11147.

60. **Zafar N**, **Mazumder R**, **Seto D**. 2002. CoreGenes: A computational tool for identifying and cataloging ``core'' genes in a set of small genomes. BMC Bioinformatics **3**.

61. **Turner D**, **Reynolds D**, **Seto D**, **Mahadevan P**. 2013. CoreGenes3.5: a webserver for the determination of core genes from sets of viral and small bacterial genomes. BMC Res Notes **6**:140.

62. **Kakoschke T**, **Kakoschke S**, **Magistro G**, **Schubert S**, **Borath M**, **Heesemann J**, **Rossier O**. 2014. The RNA chaperone Hfq impacts growth, metabolism and production of virulence factors in *Yersinia enterocolitica*. PLoS One **9**:e86113.

63. **Łoś M**, **Węgrzyn G**. 2012. Pseudolysogeny. Adv Virus Res **82**:339–49.

64. **Marusich EI**, **Kurochkina LP**, **Mesyanzhinov V V**. 1998. Chaperones in bacteriophage T4 assembly. Biochemistry Biokhimiia **63**:399–406.

65. **Matsui T**, **Griniuviene B**, **Goldberg E**, **Tsugita A**, **Tanaka N**, **Arisaka F**. 1997. Isolation and characterization of a molecular chaperone, Gp57A, of

bacteriophage T4. J Bacteriol **179**:1846–1851.

66. **Garcia-Doval C**, **van Raaij MJ**. 2012. Structure of the receptor-binding carboxy-terminal domain of bacteriophage T7 tail fibers. Proc Natl Acad Sci USA **109**:9390–9395.

67. **Cerritelli ME**, **Wall JS**, **Simon MN**, **Conway JF**, **Steven AC**. 1996. Stoichiometry and domain organization of the long tail-fiber of bacteriophage T4: a hinged viral adhesin. J Mol Biol **260**:767–780.

68. **Hashemolhosseini S**, **Stierhof YD**, **Hindennach I**, **Henning U**. 1996. Characterization of the helper proteins for the assembly of tail fibers of coliphages T4 and lambda. J Bacteriol **178**:6258–6265.

69. **Edgar R**, **Rokney A**, **Feeney M**, **Semsey S**, **Kessel M**, **Goldberg MB**, **Adhya S**, **Oppenheim AB**. 2008. Bacteriophage infection is targeted to cellular poles. Mol Microbiol **68**:1107–16.

70. **Guzev K V**, **Isaeva MP**, **Novikova OD**, **Solov'eva TF**, **Rasskazov VA**. 2005. Molecular characteristics of OmpF-like porins from pathogenic *Yersinia*. Biochemistry Biokhimiia **70**:1104–1110.

71. **Stenkova AM**, **Isaeva MP**, **Shubin FN**, **Rasskazov VA**, **Rakin A V**. 2011. Trends of the major porin gene (*ompF*) evolution: insight from the genus *Yersinia*. PLoS One **6**:e20546.

72. **Achouak W**, **Heulin T**, **PagÃ¨s J-M**. 2001. Multiple facets of bacterial porins. FEMS Microbiol Lett **199**:1–7.

73. **Silverman JA**, **Benson SA**. 1987. Bacteriophage K20 requires both the OmpF porin and lipopolysaccharide for receptor function. J Bacteriol **169**:4830–3.

74. **Traurig M**, **Misra R**. 1999. Identification of bacteriophage K20 binding regions of OmpF and lipopolysaccharide in *Escherichia coli* K-12. FEMS Microbiol Lett **181**:101–8.

75. **Gu W**, **Wang X**, **Qiu H**, **Luo X**, **Xiao D**, **Xiao Y**, **Tang L**, **Kan B**, **Jing H**. 2012. Comparative antigenic proteins and proteomics of pathogenic *Yersinia enterocolitica* bio-serotypes 1B/O: 8 and 2/O: 9 cultured at 25°C and 37°C. Microbiol Immunol **56**:583–94.

76. **Riede I**. 1987. Receptor specificity of the short tail fibres (Gp12) of T-even type *Escherichia coli* phages. Mol Gen Genet **206**:110–5.

77. **Thomassen E**, **Gielen G**, **Schütz M**, **Schoehn G**, **Abrahams JP**, **Miller S**, **van Raaij MJ**. 2003. The structure of the receptor-binding domain of the bacteriophage T4 short tail fibre reveals a knitted trimeric metal-binding fold. J Mol Biol **331**:361–73.

78. **Orquera S**, **Gölz G**, **Hertwig S**, **Hammerl J**, **Sparborth D**, **Joldic A**, **Alter T**.

2012. Control of *Campylobacter* spp. and *Yersinia enterocolitica* by virulent bacteriophages. J Mol Genet Med **6**:273–8.

79. **Noszczyńska M**, **Kasperkiewicz K**, **Duda KA**, **Podhorodecka J**, **Rabsztyn K**, **Gwizdała K**, **Świerzko AS**, **Radziejewska-Lebrecht J**, **Holst O**, **Skurnik M**. 2015. Serological characterization of the enterobacterial common antigen substitution of the lipopolysaccharide of *Yersinia enterocolitica* O:3. Microbiology **161**:219–27.

80. **Kay BA**, **Wachsmuth K**, **Gemski P**, **Feeley JC**, **Quan TJ**, **Brenner DJ**. 1983. Virulence and phenotypic characterization of *Yersinia enterocolitica* isolated from humans in the United States. J Clin Microbiol **17**:128–138.

81. **Portnoy DA**, **Falkow S**. 1981. Virulence-associated plasmids from *Yersinia enterocolitica* and *Yersinia* pestis. J Bacteriol **148**:877–883.

82. **Michiels T**, **Cornelis GR**. 1991. Secretion of hybrid proteins by the *Yersinia* Yop export system. J Bacteriol **173**:1677–1685.

83. **Demarre G**, **Guérout A-M**, **Matsumoto-Mashimo C**, **Rowe-Magnus DA**, **Marlière P**, **Mazel D**. 2005. A new family of mobilizable suicide plasmids based on broad host range R388 plasmid (IncW) and RP4 plasmid (IncPalpha) conjugative machineries and their cognate *Escherichia coli* host strains. Res Microbiol **156**:245–55.

84. **Hu B**, **Margolin W**, **Molineux IJ**, **Liu J**. 2015. Structural remodeling of bacteriophage T4 and host membranes during infection initiation. Proc Natl Acad Sci USA **112**:E4919–28.

85. **Dereeper A**, **Guignon V**, **Blanc G**, **Audic S**, **Buffet S**, **Chevenet F**, **Dufayard J-F**, **Guindon S**, **Lefort V**, **Lescot M**, **Claverie J-M**, **Gascuel O**. 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res **36**:W465–9.

86. **Kasperkiewicz K**, **Swierzko AS**, **Bartlomiejczyk MA**, **Cedzynski M**, **Noszczynska M**, **Duda KA**, **Michalski M**, **Skurnik M**. 2015. Interaction of human mannose-binding lectin (MBL) with *Yersinia enterocolitica* lipopolysaccharide. Int J Med Microbiol **305**:544–52.

Table 1. Bacterial strains, plasmids and bacteriophages

Strain	Comments	Reference
Y. enterocolitica		
6471/76 (YeO3)	Serotype O:3, wild type. Human stool isolate	(23)
6471/76-c (YeO3-c)	Virulence-plasmid cured derivative of YeO3	(23)
YeO3-ϕR1-RT-R2	ϕR1-RT resistant spontaneous derivative of YeO3	This work
YeO3-ϕR1-RT-R7	ϕR1-RT resistant spontaneous derivative of YeO3	This work
YeO3-ϕR1-RT-R9	ϕR1-RT resistant spontaneous derivative of YeO3	This work
YeO3-R1	(=YeO3-c-R1) Spontaneous rough strain	(22)
YeO3-*hfq::*Km	*Hfq::*Km-GenBlock, Km^R. Urease-negative	(Leskinen et al., submitted for publication)
YeO3-R1-Cat17	*ompF::Cat-Mu* derivative of YeO3-R1	This work
YeO3-R1-Cat17::pSW25T_OmpF	*Cis* complemented *ompF::Cat-Mu* strain	This work
YeO3-R1-M164	*waaF::Cat-Mu*. derivative of YeO3-R1. Clm^R	(79)
YeO3-R1-M196	*galU::Cat-Mu* derivative of YeO3-R1. Clm^R	(79)
YeO3-R1-M205	*hldE::Cat-Mu* derivative of YeO3-R1. Clm^R	(79)
YeO3-c-OC		(25)
YeO3-c-OCR		(25)
K14	Serotype O:9	
gc815-73	Serotype O:5,27	(80)
8081	Serotype O:8	(81)
Escherichia coli		
BL21 Star™ (DE3) PLysS		Invitrogen
DH10B		
Plasmids		
pTM100		(82)
pTM100_OmpF	Complementation plasmid with wild type *ompF* gene	This work

	cloned into pTM100	
pSW25T	Suicide vector	(83)
pSW25T_OmpF	Complementation suicide plasmid with wild type *ompF* gene cloned into pSW25T	This work
pCDF Duet-1	pCloDF13 replicon, T7lac promoter and 2 MCS sites each with an optional N-terminal His_6 tag sequence. Streptomycin resistance marker.	Novagen
pET21a(+)	ColE1 (pBR322) replicon, T7lac promoter, N-terminal T7 tag sequence and optional C-terminal His_6 Tag sequence. Ampicillin resistance marker.	Novagen
pCDF Duet-1 Gp37	Phage TG1 *ORF250* (4-1,830 bp) cloned in frame into MCS1 of pCDF Duet-1 for expression of N-terminal His_6 tagged protein Gp37.	This study
pCDF Duet-1 Gp37-Gp38	Phage TG1 *ORF251* cloned into MCS2 of pCDF Duet-1 Gp37 for co-expression of N-terminal His_6 tagged Gp37 and tail fiber assembly chaperone Gp38.	This study
pET21a(+) Gp57A	Phage TG1 *ORF143* cloned into MCS of pET21a(+) for expression of general trimerization chaperone Gp57A.	This study
Bacteriophages		
TG1		This study
ϕR1-RT		This study

Table 2. Lytic activities of phages TG1 and ϕR1-RT. The sensitivity was tested on 160 *Yersinia* species strains (Table S1) at 25°C.

***Yersinia* species**	**Phage sensitive serotypes[a]**	**Serotypes with phage sensitive (S) and resistant (R) strains**	**Phage resistant serotypes[b]**
Y. enterocolitica	O:1[2], O:2 [2], O:3 [16], O:5 [9], O5,27 [10], O:6 [2], O:7,8 [2], O:9 [13],	O:6,30 [1S/2R], O:6,31 [1S/1R]	O:1,2,3 [1], O:4 [1], O:4,32 [1], O:8 [14], O:10 [4], O:13 [1], O:13a,13b [1], O:13,7 [2], O13,18 [1], O:14 [1], O:20 [2], O:21 [3], O:25 [1], O:25,26,44 [1], O:26,44 [1], O28,50 [1], O:34 [1], O:35,36 [1], O35,52 [1], O:41(27),K1 [1], O41(27),42 [1], O:41(27),42,K1 [1], O:41,43 [1], O:41(27),43 [2], O:50 [1], K1 NT[2], NT[3]
Y. aleksiciae			O:16 [2]
Y. aldovae			UT [2]
Y. bercovieri			O:58,16 [2], NT [1], UT[2]
Y. frederiksenii			O:3 [1], O:16 [1], O:35 [1], O:48 [1], K1 NT[1], NT[1], UT[2]
Y. intermedia			O16,21 [1], O:52,54 [1], UT[2]
Y. kristensenii			O:3 [1], O:12,25 [1], NT[2], UT[3]
Y. mollareti			O:3 [1], O:59(20,36,7) [1], UT[2]
Y. nurmii			UT[1]
Y. pekkanenii			UT[1]
Y. pseudotuberculosis			I [2], O:1b [2], O:3 [2]
Y. rohdei			UT[2]
Y. ruckeri			NT[1], UT[5]

[a]The number of strains studied is given in brackets. Phage ϕR1-RT sensitivity was tested only with the 109 UH-source strains (Table S1).

[b]NT, non-typeable and either cross-reacting or not agglutinating with *Y. enterocolitica* O:3, O:5, O:8 or O:9 antisera. UT, untyped.

FIGURE LEGENDS

Figure 1. Bacteriophage ϕR1-RT and TG1 morphology by electron microscopy. **Panel A.** ϕR1-RT virions at 39,440x magnification. The virion head and tail are indicated, as well as long tail fibers (LTF) and a baseplate with protruding tail pins (B). Scale bar, 100 nm. **Panel B.** A ϕR1-RT virion at 84,320x magnification. A baseplate with protruding tail pins (B), and a neck and collar with neck fibers (A) can be observed. Scale bar, 50 nm. **Panel C.** A ϕR1-RT virion at 108,800x magnification. Suggested long tail fibers (LTF) can be seen bent up towards the head along the tail sheat, as described for bacteriophage T4 (84). Scale bar, 50 nm. **Panel D.** Bacteriophage TG1 virion shown at 150,000x magnification. A neck and collar with neck fibers (A), a baseplate with protruding tail pins (B), and an extended long tail fiber (LTF) can be observed. Scale bar indicates size in nm.

Figure 2. ProgressiveMauve alignment of phage TG1 and ϕR1-RT. The genome of ϕR1-RT [HE956709] is indicated on the top and that of TG1 [KP202158] is shown in the bottom of the figure. The degree of sequence similarity between regions is given by a similarity plot within the coloured blocks with the height of the plot proportional to the average nucleotide identity. Below these are illustrated the phage genes as outlined boxes on the plus (above horizontal) and minus (below horizontal) strands.

Figure 3. Phylogenetic analysis of the large terminase subunit protein sequences of phages TG1, ϕR1-RT and related bacteriophages. The phylogenetic analysis was constructed using "one click" at phylogeny.fr using MUSCLE for multiple alignment and PhyML for phylogeny (85).

Figure 4. Phage ϕR1-RT does not propagate at 37°C. **Panel A.** The growth curves of *Y. enterocolitica* infected with ϕR1-RT. Bacteria were cultured with different concentrations of phage particles in LB at 4°C, 10°C, 16°C, 22°C, and 37°C. Each graph represents the average of five replicates. Note the different scales used for the X-axis at different temperatures. **Panel B**. Analysis of the *ompF* gene expression (left) and protein abundance (right) at different temperatures. The mean expression levels of the *ompF* gene were obtained from RNA-sequencing analysis. The production levels of the OmpF protein was obtained from normalized mean spectral values for the proteins detected by the LC-MS/MS analysis. Error bars represent the calculated standard deviation.

Figure 5. Phages ϕR1-RT and TG1 use OmpF and LPS inner core heptose region of *Y. enterocolitica* O:3 as receptors. **Panel A.** Adsorption experiments were performed with different LPS and *ompF* mutants, with the complemented strains, and with the pseudolysogen. All strains are OmpF positive with the exception of YeO3-c-R1-Cat17. The TG1 and ϕR1-RT adsorptions to the bacteria at 5 min are shown as residual PFU percentages. Error bars indicate standard deviations. The no bacteria control (LB) and strains used for adsorptions are indicated below the columns. The LPS chemotypes of the strains are indicated on top of the columns. **Panel B.** The schematic structures of the *Y. enterocolitica* O:3 LPS molecules of different chemotypes (86). Please note that *Y. enterocolitica* O:3 carries simultaneously the S and Ra type LPS molecules. This is indicated in panel A by a plus sign. O-ag, O-antigen or O polysaccharide; OC, outer core hexasaccharide; IC, inner core; LA, lipid A.

Figure 6. Confocal immunofluorescence microscopy images of *Y. enterocolitica* cells after incubation with LTF protein Gp37 derived from phage TG1. Gp37 decorates the cell surface of *Y. enterocolitica* strain K14 of serotype O:9 (a), *Y. enterocolitica* strain gc815-73 of serotype O:5,27 (c), and *Y. enterocolitica* strain 6471/76 of serotype O:3 (e) grown at 25°C, whereas *Y. enterocolitica* strain 8081 of serotype O:8 (g) does not show cell decoration with Gp37. Similar

images to that presented in g were observed when the same strains were grown at 37°C. Differential interference contrast microscopy images of a, c, e and g, are shown in b, d, f and h, respectively. Scale bar represent size in μm.

SUPPLEMENTARY FIGURE LEGENDS

Figure S1. Plasmid constructs for chaperone-assisted expression of the distal long tail fiber protein Gp37 of phage TG1. The location of the origins of replication, antibiotic resistance genes (Sm R, streptomycin resistance; Amp R, ampicillin resistance), relevant promoters (T7 lac promoter), Lac I repressor, Multiple cloning sites (yellow) and sequences coding phage Gp37 (red), and chaperones Gp38 and Gp57A (green) are presented.

Figure S2. Map of the phage ϕR1-RT genome [HE956709]. The genes are shown by different-colored arrows. The arrow direction indicates the coding direction of the genes. The genes encoding putative proteins with an assigned function are shown in black (see also Table S3). The locations of tRNA encoding genes are shown in blue. Hypothetical proteins with an unknown function are depicted in yellow. Homing endonuclease genes are shown in orange. Putative host promoters are shown as pink triangles above the sequence, and putative rho independent terminators are shown as green triangles below the sequence.

Figure S3. Map of the phage TG1 genome [KP202158]. The genes are shown by different-colored arrows. The arrow direction indicates the coding direction of the genes. The genes encoding putative proteins with an assigned function are shown in black (see also Table S2). The locations of tRNA encoding genes are shown in blue. Hypothetical proteins with an unknown function are depicted in yellow. Homing endonuclease genes are shown in orange. Putative host promoters are shown as pink triangles above the sequence, and putative rho independent terminators are shown as green triangles below the sequence.

Figure S4. Multiple sequence alignment of the long tail fiber (Gp37) sequences of phage TG1 and ϕR1-RT. Multiple sequence alignment was performed using Clustal W via Geneious R9 software version 9.0.2. (Biomatters Ltd). Positions which have a single, fully conserved aa residue (100% similarity) are highlighted in black; aa present in 2 of the sequences are highlighted in grey. The homologous phage T4 Gp37 sequence was included in the alignment or comparison.

Figure S5. Multiple sequence alignment of the short tail fiber (Gp12) protein sequences of phage TG1 and ϕR1-RT. Multiple sequence alignment was performed using Clustal W via Geneious R9 software version 9.0.2. (Biomatters Ltd). Positions which have a single, fully conserved aa residue (100% similarity) are highlighted in black; aa present in 2 of the sequences are highlighted in grey. The homologous phage T4 Gp12 sequence was included in the sequence alignment for comparison.

Figure S6. Protein mass fingerprinting of phage TG1 Gp37. The amino acid sequence of phage TG1 Gp37 is shown. Peptide fragments from the analysis of gel slices corresponding to the reduced form of the protein and identified via protein mass fingerprinting are underlined and shown in bold.

Figure S7. Whole genome phylogeny of phages TG1, ϕR1-RT and related bacteriophages. The phylogenetic tree was generated using using “one click” at phylogeny.fr using MUSCLE for multiple alignment and PhyML for phylogeny (85).

Figure S8. Expression of N-terminal His_6 tagged phage TG1 Gp37. 4-15% Tris-HCl SDS-PAGE run at 4°C, 100V. Lane 1, pCDF-Duet-1 Gp37 (unheated sample); Lane 2, pCDF-Duet-1 Gp37 under reduced conditions (sample heated at 100°C for 10 min in the presence of SDS and β-mercaptoethanol); Lane 3, pCDF-Duet-1 Gp37-Gp38 (unheated sample); Lane 4, pCDF-Duet-1 Gp37-Gp38 reduced conditions; Lane 5, pCDF-Duet-1 Gp37-Gp38 co-expressed with pET21a(+) Gp57A (unheated sample); Lane 6 pCDF-Duet-1 Gp37-Gp38 co-expressed with pET21a(+) Gp57A reduced conditions; Lane 7, pCDF-Duet-1 Gp37 co-expressed with pET21a(+) Gp57A (unheated sample); Lane 8, pCDF-Duet-1 Ggp37 co-expressed with pET21a(+) Gp57A reduced conditions; Lane M, molecular weight markers.

Figure S9. CLUSTAL W multiple sequence alignment of the OmpF proteins of *Y. enterocolitica* and related species. The external loops, indicated by brown highlighting and box, were identified based on OmpF alignment of YeO3 OmpF with that of *E. coli*, shown at the bottom. The N-terminal signal-peptide is indicated by blue highlighting and box. Use the zoom-in option to see details of the alignment.

SUPPLEMENTARY TABLES

Table S1. List of *Yersinia* strains used in phage host range experiments

Table S2. Primers used in this work

Table S3. Bacterial strains used to examine the cross infectivity of phage TG1

Table S4. Annotations of bacteriophage TG1 and ϕR1-RT genes

Table S5. Putative Phage TG1 promoters

Table S6. Putative Phage ϕR1-RT promoters

Table S7. Predicted terminator sequences of phage TG1

Table S8. Predicted terminator sequences of phage ϕR1-RT

Table S9. TG1 genome BLASTN analysis

Table S10. Whole genome sequencing statistics after de novo assembly of ϕR1-RT resistant mutants.

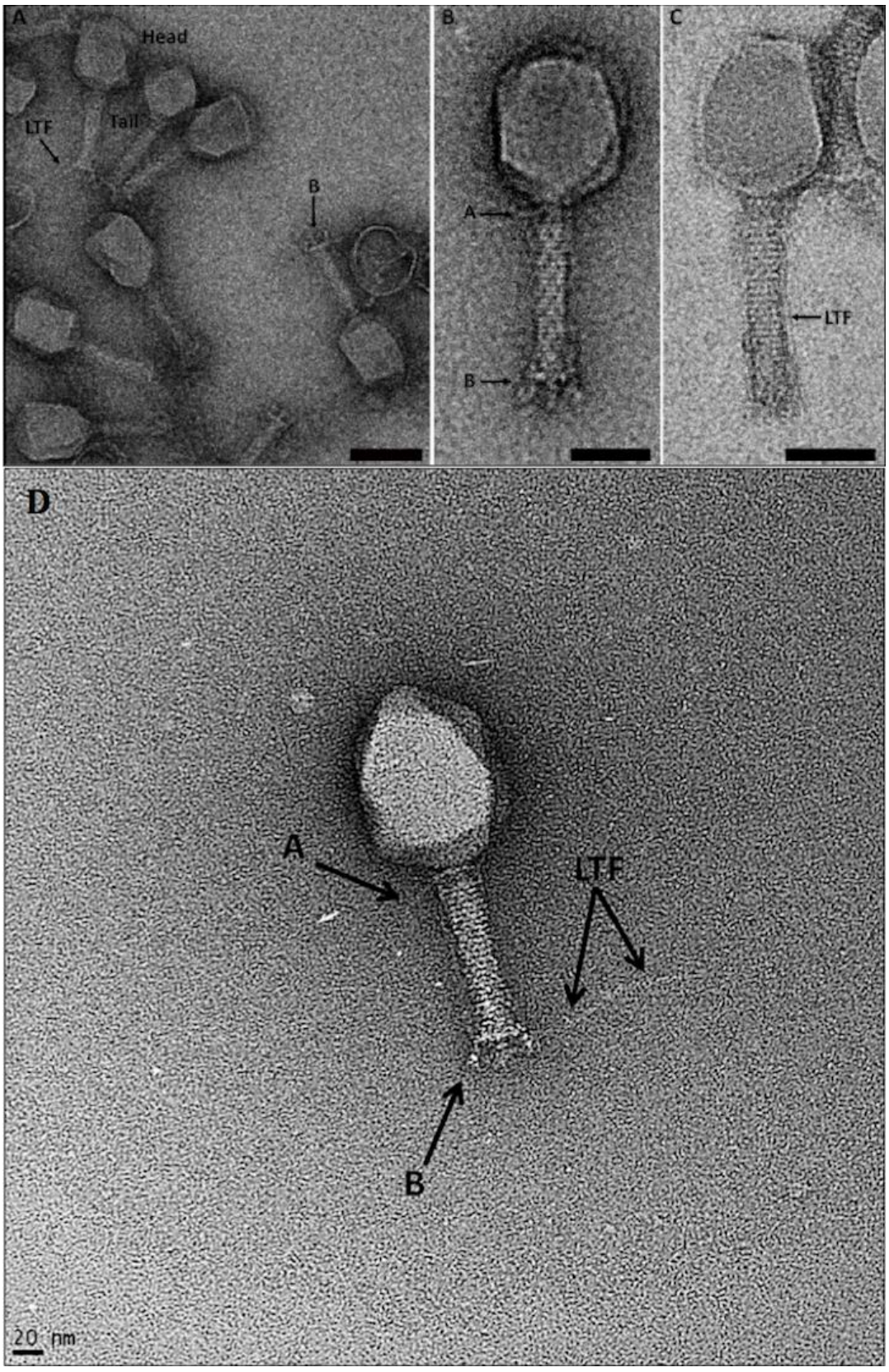

Figure 1.

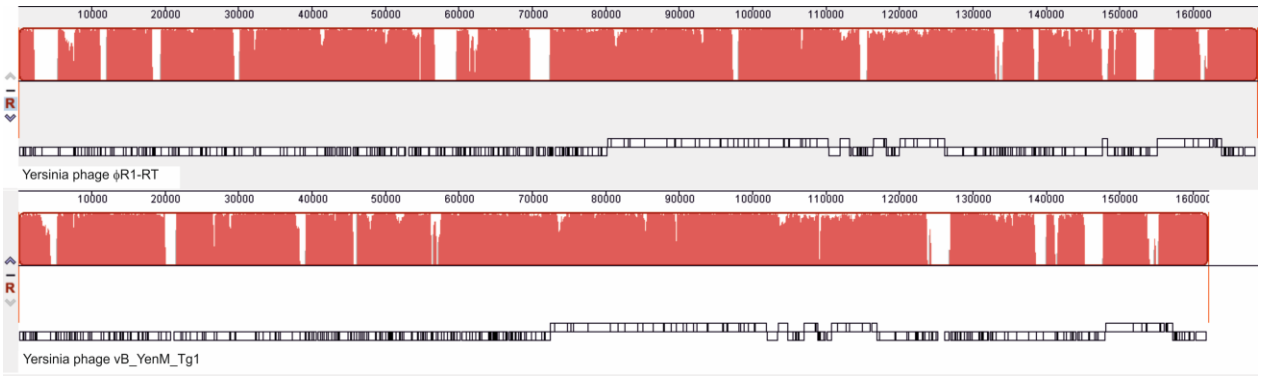
10000
20000
30000
40000
50000
60000
70000
80000
90000
100000
110000
120000
130000
140000
150000
160000
Yersinia phage ϕR1-RT
Yersinia phage vB_YenM_Tg1

Figure 2.

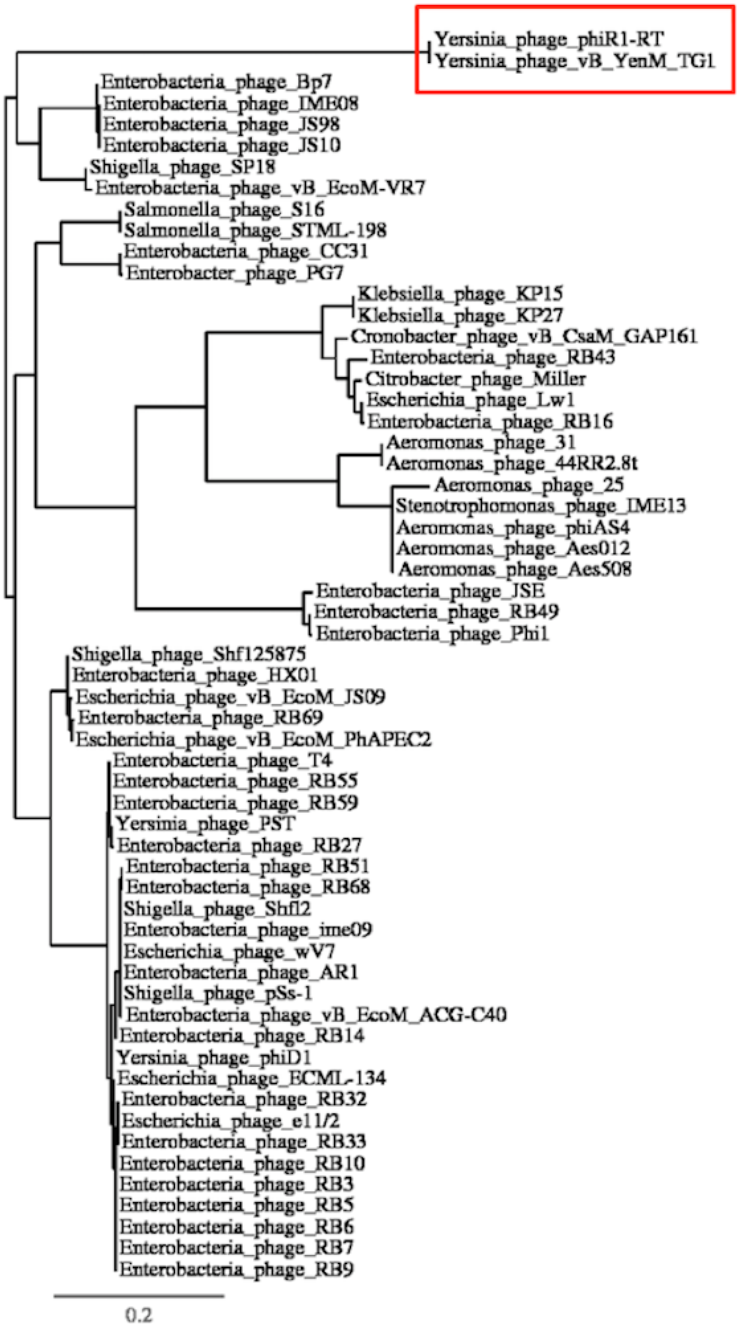
Yersinia_phage_phiR1-RT
Yersinia_phage_vB_YenM_TG1
Enterobacteria_phage_Bp7
Enterobacteria_phage_IME08
Enterobacteria_phage_JS98
Enterobacteria_phage_JS10
Shigella_phage_SP18
Enterobacteria_phage_vB_EcoM-VR7
Salmonella_phage_S16
Salmonella_phage_STML-198
Enterobacteria_phage_CC31
Enterobacter_phage_PG7
Klebsiella_phage_KP15
Klebsiella_phage_KP27
Cronobacter_phage_vB_CsaM_GAP161
Enterobacteria_phage_RB43
Citrobacter_phage_Miller
Escherichia_phage_Lw1
Enterobacteria_phage_RB16
Aeromonas_phage_31
Aeromonas_phage_44RR2.8t
Aeromonas_phage_25
Stenotrophomonas_phage_IME13
Aeromonas_phage_phiAS4
Aeromonas_phage_Aes012
Aeromonas_phage_Aes508
Enterobacteria_phage_JSE
Enterobacteria_phage_RB49
Enterobacteria_phage_Phi1
Shigella_phage_Shf125875
Enterobacteria_phage_HX01
Escherichia_phage_vB_EcoM_JS09
Enterobacteria_phage_RB69
Escherichia_phage_vB_EcoM_PhAPEC2
Enterobacteria_phage_T4
Enterobacteria_phage_RB55
Enterobacteria_phage_RB59
Yersinia_phage_PST
Enterobacteria_phage_RB27
Enterobacteria_phage_RB51
Enterobacteria_phage_RB68
Shigella_phage_Shf12
Enterobacteria_phage_ime09
Escherichia_phage_wV7
Enterobacteria_phage_AR1
Shigella_phage_pSs-1
Enterobacteria_phage_vB_EcoM_ACG-C40
Enterobacteria_phage_RB14
Yersinia_phage_phiD1
Escherichia_phage_ECML-134
Enterobacteria_phage_RB32
Escherichia_phage_e11/2
Enterobacteria_phage_RB33
Enterobacteria_phage_RB10
Enterobacteria_phage_RB3
Enterobacteria_phage_RB5
Enterobacteria_phage_RB6
Enterobacteria_phage_RB7
Enterobacteria_phage_RB9
0.2

Figure 3

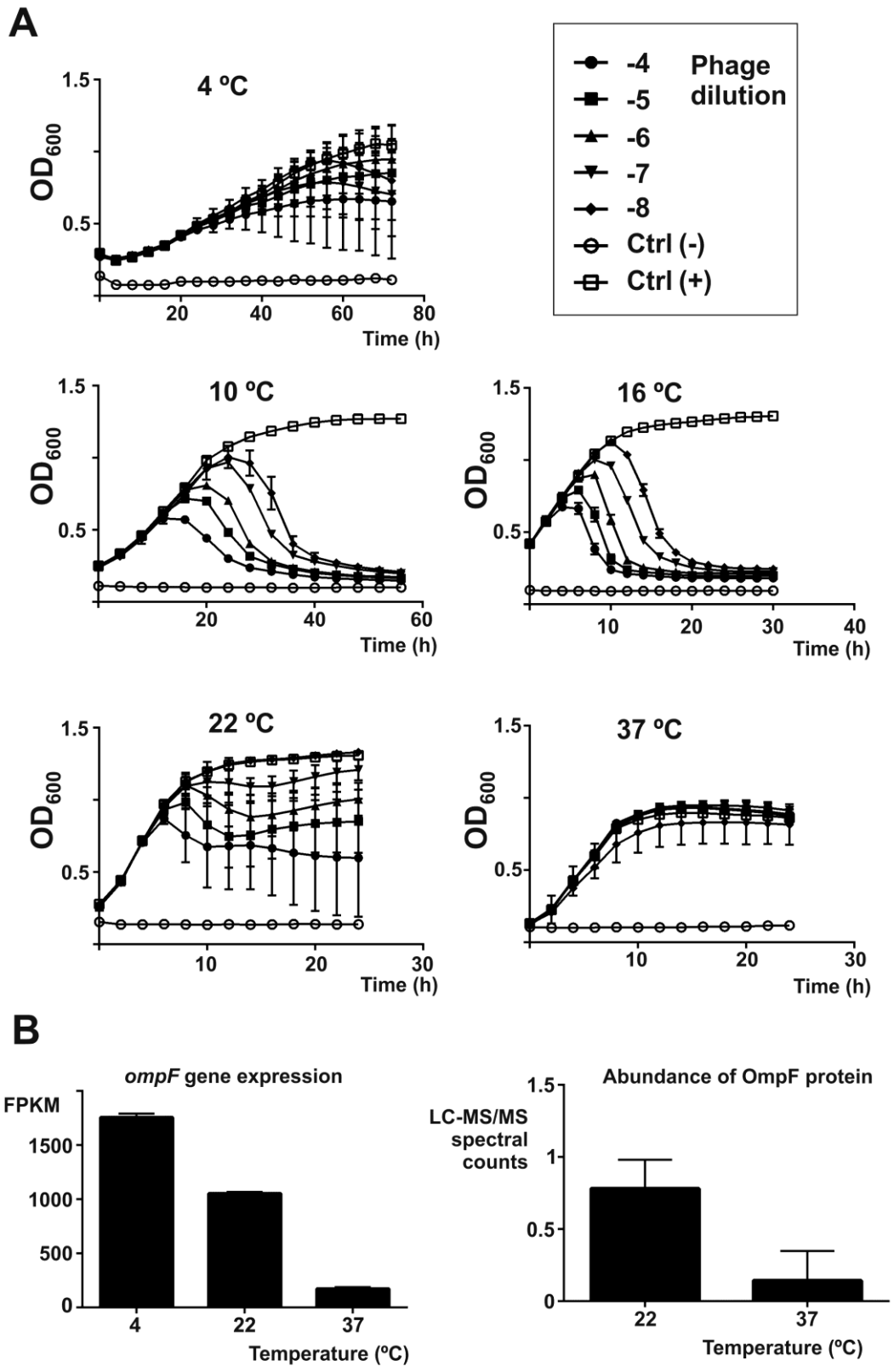
A
4 ºC
10 ºC
16 ºC
22 ºC
37 ºC
OD600
Time (h)
Phage dilution
-4
-5
-6
-7
-8
Ctrl (-)
Ctrl (+)
B
ompF gene expression
FPKM
Temperature (ºC)
Abundance of OmpF protein
LC-MS/MS spectral counts
Temperature (ºC)

Figure 4

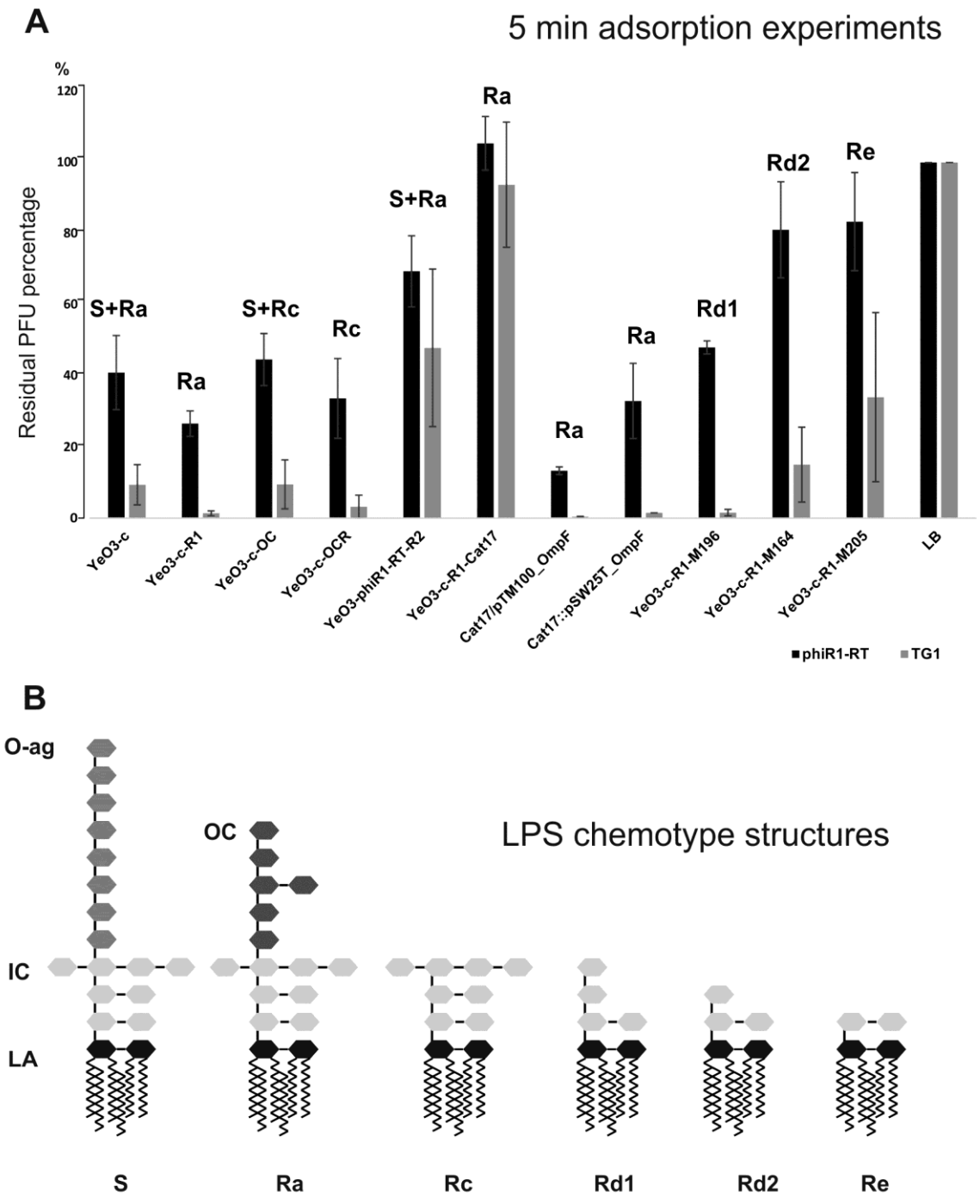
A
5 min adsorption experiments
%
Residual PFU percentage
120
100
80
60
40
20
0
S+Ra
Ra
S+Rc
Rc
S+Ra
Ra
Ra
Rd1
Ra
Rd2
Re
YeO3-c
Yeo3-c-R1
YeO3-c-OC
YeO3-c-OCR
YeO3-phiR1-RT-R2
YeO3-c-R1-Cat17
Cat17/pTM100_OmpF
Cat17::pSW25T_OmpF
YeO3-c-R1-M196
YeO3-c-R1-M164
YeO3-c-R1-M205
LB
phiR1-RT
TG1
B
O-ag
OC
LPS chemotype structures
IC
LA
S
Ra
Rc
Rd1
Rd2
Re

Figure 5

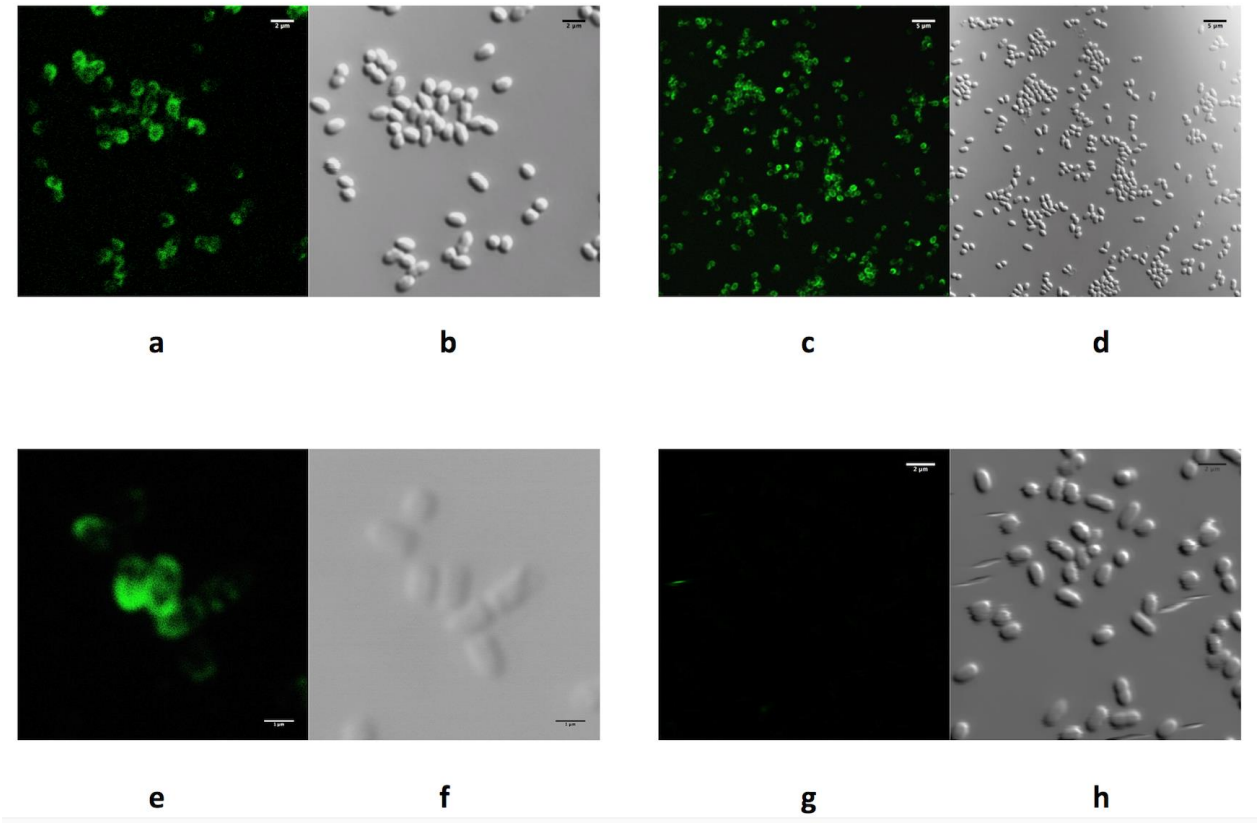

Figure 6